

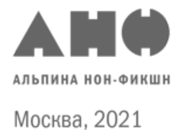
МАЙКЛ ГРАЦИАНО
НАУКА СОЗНАНИЯ
СОВРЕМЕННАЯ ТЕОРИЯ СУБЪЕКТИВНОГО ОПЫТА
П
КНИГИ ПОЛИТЕХА

майкл грациано

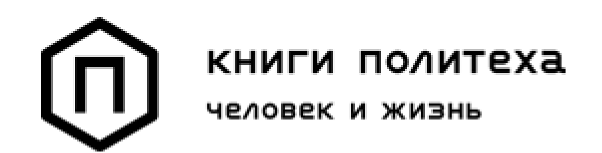
наука сознания

современная теория субъективного опыта

перевод с английского



Москва, 2021


книги политеха
человек и жизнь



“КНИГИ ПОЛИТЕХА” — партнерский проект ПОЛИТЕХНИЧЕСКОГО МУЗЕЯ, издательств CORPUS, “АЛЬПИНА НОН-ФИКШН” и “БОМБОРА”.

В серии выходят лучшие современные и классические книги о науке и технологиях — все они отобраны и проверены учеными и отраслевыми специалистами.

Серия “Книги Политеха” — это пять коллекций, связанных с темами постоянной экспозиции Политехнического музея:

“Человек и жизнь” — мир живого, от устройства мозга до биотехнологий.

“Цифры и алгоритмы” — математика, искусственный интеллект и цифровые технологии.

“Земля и Вселенная” — происхождение мира, небесные тела, освоение космоса, науки о Земле.

“Материя и материалы” — устройство мира с точки зрения физики и химии.

“Идеи и технологии” — наука и технологии, их прошлое и будущее.



Политехнический музей представляет новый взгляд на экспозицию, посвященную науке и технологиям. Спустя столетие для музея вновь становятся важными

мысль и идея, а не предмет, ими созданный.

Научная часть постоянной экспозиции впервые визуализирует устройство мира с точки зрения современной науки — от орбиталей электрона до черной дыры,

от структуры ДНК до нейронных сетей.

Историческая часть постоянной экспозиции рассказывает о достижениях российских инженеров и изобретателей как части мировой технологической культуры — от самоходного судна Ивана Кулибина до экспериментов по

термоядерному синтезу и компьютера на основе троичной логики.

Политех делает все, чтобы встреча человека и науки состоялась. Чтобы наука осталась в жизни человека навсегда. Чтобы просвещение стало нашим

общим будущим.

Подробнее о Политехе и его проектах — на **polytech.one**

Бену и Эли

Глава

1

Слон в комнате[1]

Когда моему сыну было три года, я “озвучивал” его любимого плюшевого слоника, делая вид, что тот умеет говорить. В столь юном возрасте ребенок еще не мог понять, насколько никудышный из меня чревовещатель, так что это срабатывало. Сын был в восторге. В последующие несколько лет я оттачивал свою технику и удивлялся загадочной силе этой иллюзии. Чревовещание — не просто голос, исходящий из игрушки, как если бы в ней был спрятан динамик. Даже голосом бездарного исполнителя вроде меня творится волшебство: игрушка обретает жизнь, личность, она будто бы наполняется сознанием.

Очевидно, в человеческом мозге должны быть механизмы, вынуждающие приписывать игрушке сознание. Но они развились не для того, чтобы мы развлекались чревовещанием. Человек — общественное животное, и он регулярно применяет эти механизмы по отношению к другим. Когда я с кем-то разговариваю, у меня автоматически возникает представление о мыслях, эмоциях и переживаниях собеседника. Конечно, я не вижу его психическую деятельность напрямую. Но мой мозг строит удобную модель психической деятельности и проецирует ее на

этого человека, т.е. поступает так же, как мой сын с “заговорившей” игрушкой.

Мы обращаемся подобным образом не только с людьми. Мы приписываем сознание домашним кошкам и собакам, а некоторые готовы поклясться, что у них и растения в горшках обладают сознанием. Древние люди были убеждены, что разумны деревья и реки; дети видят личность в своих любимых игрушках, и, черт возьми, вчера я наорал на свой компьютер! Но речь, собственно, не идет о том, чтобы интеллектуальными усилиями выяснить, есть ли у чего-то психика, или прийти к заключению, что могло бы быть в этом психическом мире, — хотя это мы тоже делаем. Я говорю об интуитивном чутье, которое часто подводит нас, но иногда очень убедительно сообщает нам, что некий предмет прямо-таки излучает сознание.

Размышляя о чревовещании, я стал задумываться о том, что и мое собственное сознание, и сознание, приписываемое другим, могут исходить из одного и того же источника. Возможно, есть объединяющее объяснение: мы автоматически строим модели психики и проецируем их как на самих себя, так и на других людей. Наши интуитивные представления о загадочной сознающей личности, наше убеждение, что она есть во мне, в тебе, в домашнем любимце или игрушке, — могут быть основаны на упрощенных, но удобных моделях, комплексах информации, которые выстраивает мозг, чтобы понять свой мир.

Столь глубокие озарения приходят только к тому, кто разговаривает с плюшевыми слониками. Благодаря им я сосредоточился в своей научной работе на исследованиях сознания.

В течение 20 лет я занимался более традиционными проблемами нейронауки: как мозг отслеживает происходящее в непосредственной близости к телу и как контролирует сложные движения в этом пространстве[1]. Оказалось, что базовая подготовка в практике науки о мозге пригодилась для построения теории сознания. В 2010 г., исходя из данных нейронауки, психологии и эволюционной биологии и добавив некоторые инженерные прозрения, мы с коллегами начали набрасывать контуры того, что назвали теорией схемы внимания[2]. Эта теория — часть глобальной смены парадигмы в научном сообществе[3]. Новый подход не решает так называемую трудную проблему сознания — как физический мозг может создавать нефизические сущности[4]. Но он объясняет, почему люди могут ошибочно полагать, что трудная проблема существует в принципе, почему это ошибочное представление заложено в нас так глубоко, что вряд ли мы сможем его изменить, и почему оно выгодно (если не необходимо) для функционирования мозга.

Вначале я рассматривал эту теорию с позиции социальных взаимодействий. Но в ее глубине лежит более общее свойство мозга: знание, основанное на моделях[5]. Мозг строит внутренние модели — изменчивые сгустки информации, которые образуются постоянно и непроизвольно, подобно пузырям значений, лежащих ниже уровня высшего мышления или языка. Эти внутренние модели отражают важные для отслеживания элементы, иногда ими являются внешние объекты, а иногда — свойства самой личности. Отражения оказываются упрощенными и искаженными, будто реальность изображена импрессионистом или кубистом, а мы описываем их как самую что ни на есть

действительность. Мы не можем этого не делать, в нас это встроено. Наше интуитивное понимание окружающего мира и самих себя всегда искажено и упрощено и всегда основано на этих внутренних моделях.

В описываемой теории метафизические догадки о самих себе, о сознании как нефизической внутренней сущности (иногда ее именуют “призраком в машине”[6]) строятся на основе определенной внутренней модели. Я называю ее схемой внимания: почему — объясню дальше. Это упрощенное описание того, как мозг вбирает и обрабатывает информацию; оно представляет собой эффективный для мозга способ понимать и отслеживать свои собственные возможности. Подобного рода внутреннюю модель можно использовать, хоть и в меньшей степени, чтобы отслеживать и прогнозировать мышление и поведение других людей.

Подход, основанный на моделях, может выглядеть так, будто сознание мы со счетов сбрасываем, — но нет. Внутренняя модель, сообщающая нам о нашем сознании, богата, она обладает глубиной и протяженностью. Скорее всего, без нее не обойтись. Практически ничего из того, что мы осуществляем — будь то восприятие, мышление, активная деятельность или социальные взаимоотношения, — не работало бы без этой части системы.

В своей книге я буду пользоваться терминами “сознание”, “субъективное осознание” и “субъективный опыт (субъективное переживание)” как взаимозаменяемыми, хотя и соглашусь, что ученые не всегда применяют их в одном и том же контексте. Слово “сознание” окружено особенно широким ореолом нечетких коннотаций. Прежде чем

я доберусь до того, что имею в виду, я хочу уточнить, чего точно в виду *не имею*. Иногда люди представляют себе сознание как способность понимать, кто ты такой, понимать свой путь в жизни. Другие считают, что это скорее способность воспринимать окружающий мир, обрабатывать информацию и принимать на ее основе разумные решения. Ничего из этого я в виду не имею.

Чаще всего внутренний опыт трактуют, пожалуй, как поток сознания — постоянно меняющееся, калейдоскопическое содержание психической деятельности. То самое брожение в голове, которое Джеймс Джойс запечатлел в знаменитом романе “Улисс”[7]. Писатель скрупулезно записывал сменяющие друг друга виды, звуки, ощущения от прикосновений, вкусы и запахи, всплывающие воспоминания о недавнем и далеком прошлом, неостановимый внутренний диалог, противоборствующие эмоции и фантазии (некоторые из них сочли настолько возмутительными, что книгу сначала запретили: судебное разбирательство 1933 г. “Соединенные Штаты против книги под названием “Улисс”» обеспечило нас современным юридическим определением непристойности). Но и этого я не имею в виду, говоря о сознании. Такой поток материала не поддается четкому определению, а его научное изучение было бы непосильным хотя бы по причине его объема.

Представьте вместо этого, что мы сложили в ведро тысячу мелких предметов. Можно подобно Джойсу составить пространный список всего, что там есть. Но можно также задать вопрос более фундаментальный: как насчет самого ведра? Не будем пока о его содержимом. Из чего сделано ведро, откуда оно

взялось? Как человек вообще что-то осознает? Сознание не может быть просто информацией внутри нас, потому что в каждый конкретный момент мы осознаем лишь крошечную часть того, что содержится в мозге. Чтобы мы осознали кусочек информации, с ней что-то должно произойти. Так из-за чего же происходит осознание? Этот вопрос все больше занимал философов и ученых[8]. Термин “сознание” стал означать акт осознания чего-то, а не материал, который вы осознаете.

Я подозреваю, что постепенный сдвиг в философии от сосредоточения на элементах в потоке сознания к *акту осознания* некоторым образом связан с развитием за последние полвека компьютерных технологий. С усовершенствованием информационных технологий информационное содержание сознания перестало быть загадкой, а сам акт его осознания, переживания опыта отдалился и кажется неразрешимым. Давайте рассмотрим несколько примеров.

Вы можете подсоединить к компьютеру цифровую камеру и дать системе команду обрабатывать входящую зрительную информацию. Компьютер в состоянии определить цвета, формы, размеры объектов, распознать их. Человеческий мозг делает что-то похожее. Разница в том, что у людей еще есть субъективное переживание того, что они видят. Мы не просто регистрируем информацию о том, что предмет красный: мы переживаем *опыт* красноты. Мы что-то *чувствуем*, когда видим. Современный компьютер может обработать зрительный образ, но инженеры еще не придумали, как заставить его осознать полученную информацию.

А теперь рассмотрим кое-что более личное, нежели зрительное восприятие: ваши воспоминания, которые определяют ваш жизненный путь. Постоянно бурлящие воспоминания — типичный пример джойсовского потока сознания. И все же мы умеем строить машины, которые хранят и извлекают их. На это способен каждый компьютер, а ученые знают общие принципы, да чуть ли не все детали того, как воспоминания хранятся в мозге. Память — не какая-то принципиальная загадка. Но не является она и причиной сознания. Содержание сознания — в данном примере это воспоминания — совсем не то же самое, что акт *осознания* воспоминания.

Приведу еще один пример: принятие решений. Тайну человеческого сознания лучше всего характеризует именно эта наша способность. Мы берем информацию, обрабатываем ее, оцениваем — и совершаем выбор, что делать дальше. Но я бы все-таки сказал, что сознание не является неотъемлемой частью принятия решений. Все компьютеры делают это. В каком-то смысле это и есть задача компьютера. Он берет информацию, производит с ней какие-то действия и пользуется ею, чтобы выбрать один из многих вариантов дальнейших действий. Большинство решений в человеческом мозге — вероятно, десятки тысяч в день — принимаются автоматически, без участия субъективного переживания опыта. В некоторых особых случаях мы сообщаем о субъективном осознании принятия решения. Иногда мы называем это намерением, выбором или свободой воли. Но простая способность принять решение не требует сознания.

Эти и многие другие примеры показывают, как именно расцвет компьютерных технологий позволил увидеть различие между содержанием сознания

(которое становится все понятнее на инженерном уровне) и актом осознания этого содержания. Мне интересна вторая, более значимая часть этой головоломки: как нам вообще удается получить субъективное переживание чего бы то ни было?

Некоторые считают, что это ограничивающий подход. Меня часто спрашивают: а как же память? Как же осознанный выбор? Самопонимание? А намерения и убеждения? Разве это не основа сознания? Я согласен: все перечисленные вопросы важны, они суть главные предметы, лежащие в ведре человеческого сознания. Но в них нет принципиальной загадки. Это вопросы обработки информации, и мы можем представить себе, хотя бы в общих чертах, как они ставятся в инженерном смысле. Принципиальная загадка — само ведро. Что такое сознание — из чего оно сделано? Как в него попадают, в чем выгода в него попасть и почему в него попадает так мало из содержащегося в мозге?

Ученые традиционно полагали, что, применяя научный подход, нечто столь аморфное и ускользающее понять невозможно. Но, принимая в расчет недавние наблюдения и выводы, я практически уверен, что сознание настолько же поддается пониманию и построению, как и обработка зрительной информации, память, принятие решений или любой другой элемент из его содержания.

Я и раньше много писал о сознании. Но эта книга целиком обращена к широкой аудитории. В ней я пытаюсь как можно проще и четче разъяснить многообещающую научную теорию сознания, которая одинаково приложима и к биологическому мозгу, и к искусственной машине.

Завязка нескольких следующих глав — эволюция. Я буду описывать развитие усложнения нервной системы начиная от полумиллиарда лет назад, когда появились нейроны (одни из видов клеток, из которых состоит мозг). По пути я стану вводить элементы теории схемы внимания, и к шестой главе у нас будут готовы основные строительные леса.

Затем я обращусь к тому, как данная теория взаимодействует с другими. Это одна из полудюжины основных теорий сознания, которые сейчас набирают вес в научной литературе. Согласно моим представлениям, которые я стараюсь передать в книге, эти теории не всегда следует рассматривать как конкурирующие — и не стоит гадать, какая из них перебьет всех соперников. Несмотря на их различия (а я действительно со многим в них не согласен) между этими теориями могут также быть странные, потайные связи. В каждой есть важные мысли. Мне кажется, мы начинаем видеть на горизонте проблески общего мнения… Или, скорее, сети согласованных представлений.

В последних главах я углублюсь в потенциальные технологические последствия этой теории. Мы близки к пониманию сознания, достаточному для его конструирования. А когда у нас это получится, новые технологии, вероятно, изменят лицо нашей цивилизации. Сознательные машины — лишь первый шаг. Если сознание удастся построить, то тогда в принципе и разум можно будет переносить с одного устройства на другое. Более отдаленное, но возможное следствие — считывание данных с человеческого мозга и перенос психического мира этого человека на искусственную платформу[9]. Такие технологии могли бы позволить личности жить вечно и исследовать

враждебные биологическим телам среды — например, межзвездное пространство. У нас на пути больше не стоят законы физики, необходимо лишь изобрести нужные устройства.

Если сознание можно понять с научной и инженерной точек зрения, то данная тема перестает быть просто философским развлечением для ученых. Она становится непосредственно важной практической задачей. Дальше в книге я опишу возможные применения сознания во многих вариантах технологического будущего; некоторые из них покажутся привлекательными, а некоторые, скажем прямо, ужасными. Но, хорошо это или плохо, я практически уверен, что мы стремительно движемся к научному пониманию сознания и способности сделать его искусственно.

Глава

2

Крабы и осьминоги

Самовоспроизводящаяся бактериальная жизнь появилась на Земле примерно 4 млрд лет назад. На протяжении почти всей истории Земли жизнь оставалась на одноклеточном уровне, и ничего похожего на нервную систему не существовало вплоть до 600–700 млн лет назад. В теории схемы внимания сознание основано на определенном способе обработки информации нервной системой. Ключевой элемент этой теории (и, я полагаю, любого развитого интеллекта) — внимание: способность мозга в каждый момент времени сосредоточивать свои ограниченные ресурсы на небольшом фрагменте мира, чтобы получить бóльшую глубину обработки. В этой и нескольких следующих главах я рассмотрю, как внимание могло развиться от древних животных до людей и как вместе с ним могло появиться свойство, которое мы называем сознанием[1].

Начнем с морских губок, они “помогут” очертить границы эволюции нервной системы. Губки — самые примитивные многоклеточные организмы, у них нет так называемого плана тела, нет конечностей, нет мышц, — и нервы им не нужны. Они закрепились на дне океана и фильтруют питательные вещества подобно ситу. Но у нас есть общие с губками гены, в

том числе не менее 25 из тех, которые у людей помогают структурировать нервную систему[2]. У губок те же самые гены могут выполнять более простые функции, например участвовать в коммуникации клеток друг с другом. Губки как будто балансируют на эволюционной грани нервной системы.

Считается, что последний общий у нас с ними предок существовал в диапазоне от 700 до 600 млн лет назад (см. шкалу времени на рис. 2.1)[3].



Рис. 2.1 Описываемые в этой главе беспозвоночные и примерное время их появления

Другие древние животные — медузы — напротив, обладают нервной системой. Медузы плохо сохраняются в окаменелостях, но, анализируя их генетические взаимосвязи с другими животными, биологи предполагают, что они могли отделиться от остального животного царства примерно 650 млн лет назад[4]. Эти цифры, возможно, изменятся с получением новых данных, но в качестве правдоподобного предположения скажем, что нейроны — базовые клеточные компоненты нервной системы — впервые появились в животном царстве между губками и медузами.

Нейрон по сути своей — это клетка, передающая сигнал. Волна электрохимической энергии прокатывается по мембране клетки от одного края нейрона до другого со скоростью чуть более 60 м/с и действует на другой нейрон, мышцу или железу. Самые первые нервные системы могли быть устроены как простые сети нейронов, пронизывающие тело и

соединяющие мышцы. По этому принципу нервных сетей существуют гидры[5]. Это небольшие водные создания, прозрачные, похожие на цветы, в роли тела у них выступает мешок со множеством щупалец; они принадлежат к той же древней категории, что и медузы. Если коснуться гидры в одном месте, нервная сеть распространит сигнал повсюду и вся гидра дернется.

Нервная сеть не обрабатывает информацию — не извлекает из нее какого-то значения. Она просто передает сигналы по телу, соединяет сенсорный стимул (прикосновение) с мышечной реакцией (подергивание). Но после возникновения нервной сети нервные системы довольно быстро перешли на новый уровень сложности: речь идет о способности усиливать некоторые сигналы относительно других. Форсирование сигнала — простой, но мощный прием, один из основных способов, посредством которых нейроны манипулируют информацией. Это базовый компонент практически всех известных нам вычислений, происходящих в мозге.

Один из наиболее изученных примеров — глаз краба[6]. У этого животного сложные глаза со множеством детекторов, в каждом из которых есть нейрон. Когда свет падает на детектор, он активирует находящийся внутри нейрон. Пока все идет как надо. Но добавим щепотку сложности: каждый нейрон связан с ближайшими соседями и по этим связям они соревнуются друг с другом. Когда активируется нейрон в одном детекторе, он пытается приглушить активность нейронов в соседних, подобно человеку в толпе, который старается кричать громче всех и заглушить тех, кто рядом с ним.

В результате получается, что, если на глаз краба направлено размытое пятно света и на один из детекторов попадает самая яркая его часть, нейрон в этом детекторе развивает высокую активность, побеждает в соревновании и отключает соседей. Паттерн активности набора детекторов сигнализирует не только о пятне света, но и о том, что вокруг пятна — кольцо темноты. Таким образом, сигнал усилен. Глаз краба берет размытую реальность из оттенков серого и повышает ее резкость, получая контрастную картинку, где тени темнее, а яркое ярче. Усиление сигнала — прямое следствие того, что нейроны подавляют своих соседей: этот процесс называется *латеральным торможением*[7].

Описанный механизм в глазу краба, пожалуй, один из самых простых и базовых примеров, модельный экземпляр внимания. Сигналы соревнуются друг с другом, победители усиливаются за счет проигравших, и победившие сигналы затем влияют на движения животного. *Это и есть* моделирующая сущность внимания. Наше, человеческое, внимание — просто усложненная версия, состоящая из подобных компонентов. Латеральное торможение, такое же как в глазу у краба, можно найти на любой стадии обработки информации в нервной системе человека — от глаза до высших уровней мышления в коре головного мозга. Зарождение внимания лежит глубоко в эволюционной древности, ему более полумиллиарда лет, и произошло оно от удивительно простого нововведения (на тот момент, разумеется).

Крабы принадлежат к обширной группе животных под названием “членистоногие”, в которую входят пауки, насекомые и подобные им создания с твердыми сегментированными экзоскелетами. Они отделились

от других животных около 600 млн лет назад[8]. Самое известное вымершее членистоногое, у которого сегодня больше всего поклонников, — это трилобит, существо из сочленений и ножек, похожее на маленького мечехвоста, которое главным образом копошилось на дне кембрийских морей примерно 540 млн лет назад. Когда трилобиты вымерли и оказались погребены в тончайшей взвеси осадка на дне океана, они превратились в окаменелости, у которых во всех подробностях сохранились фасеточные глаза[9]. Если вы вглядитесь в выпученные очи ископаемого трилобита через лупу, то, скорее всего, вам удастся увидеть нетронутую мозаику отдельных детекторов. Судя по ископаемым остаткам, глаза трилобитов весьма напоминали глаза современных крабов и, должно быть, в них использовался тот же способ соревнования между соседними детекторами, чтобы повысить резкость обзора древнего морского дна.

Представьте себе животное, которое собирается по частям, сосредоточиваясь на каждом конкретном фрагменте. У такого животного любая часть тела будет работать как отдельный механизм, отбирая себе информацию и выделяя самые перцептивно значимые (насыщенные) сигналы. Один глаз скажет: “Вот самое яркое пятно, не реагируй на остальные”. А в это же время одна из ног пожалуется: “Меня только что сильно ткнули вот сюда, не обращай внимания на легкие прикосновения рядом!” Животное, способное лишь на такое, будет действовать как сборище отдельных “деятелей”, которые склеены друг с другом просто физически, при этом каждый выкрикивает свои сигналы и вызывает свои собственные действия. Поведение такого животного будет в лучшем случае беспорядочным.

Для того чтобы непротиворечиво реагировать на окружающую среду, животному нужно более централизованное внимание. Могут ли отдельные источники входящей информации — глаза, тело, ноги, уши, химические сенсоры — объединить свои данные, чтобы создать глобальную иерархию и отсортировать соревнование между сигналами? Подобное взаимодействие позволило бы животному выделить тот самый яркий объект в окружающей среде, который показался бы важнее всего в данный момент, и отреагировать единым, значимым образом.

Никто не знает, когда впервые появилось такое централизованное внимание, — в частности, потому что никто не знает точно, у каких животных оно есть, а у каких нет. У позвоночных есть центральный процессор внимания, который я опишу в следующей главе. Но у беспозвоночных механизмы внимания не так тщательно изучены. У многих видов животных, например кольчатых червей и брюхоногих моллюсков, нет централизованного мозга. У них есть кластеры нейронов, или ганглии, разбросанные по всему телу для локальной обработки информации[10]. Вероятно, нет у этих животных и централизованного внимания.

Более подходящие кандидаты на обладание им — членистоногие, такие как крабы, насекомые и пауки. У них есть центральный мозг или, по крайней мере, скопление нейронов в голове, которое обильнее всех остальных в их телах[11]. Эти крупные ганглии могли развиться в том числе из-за каких-то потребностей зрения. Поскольку глаза расположены в голове, а зрение — самое сложное и нагруженное информацией чувство, голова получает самую большую долю нейронов. Некоторые аспекты обоняния, вкуса, слуха

и осязания также сходятся в этом центральном ганглии. Насекомые мозговитее, чем мы думаем. Когда вы пытаетесь прихлопнуть муху, а ей практически всегда удается ускользнуть — это не просто рефлекс. Скорее, у мухи есть то, что мы называем централизованным вниманием — способность быстро сосредоточить ресурсы обработки информации на том фрагменте окружающего мира, который важнее всего в данный момент, чтобы выдать скоординированную реакцию[12].

Осьминоги — суперзвезды среди беспозвоночных: их интеллект поразителен. Их относят к моллюскам — как улиток и мидий. Моллюски появились, вероятно, около 550 млн лет назад и оставались довольно просто организованными — по крайней мере, в том, что касается нервной системы, — на протяжении сотен миллионов лет[13]. У одной из ветвей развития, головоногих моллюсков, постепенно развились сложный мозг и сложное поведение; формой они стали напоминать современных осьминогов примерно 300 млн лет назад[14].

Осьминоги, кальмары и каракатицы — поистине инопланетяне по отношению к нам[15]. Так далеко от нас на древе жизни нет других разумных животных. Они показывают нам, что мозговитый ум — не единичный феномен, так как он независимо развивался как минимум дважды: один раз в случае позвоночных, а затем снова у беспозвоночных. Осьминоги прекрасные хищники, а полагаются они на зрение. Хороший хищник должен обладать лучшей координацией и умом, чем его добыча, а использование зрения, чтобы обнаружить и распознать жертву, требует особо крупных моделирующих мощностей. Ни у какой другой

сенсорной системы нет подобного пожарного шланга, хлещущего внутрь всевозможной информацией, и нет подобной необходимости в грамотном способе сосредоточиваться на полезных фрагментах этой информации. А значит, внимание для такого хищника решает всё. Может быть, этот-то образ жизни осьминога и повлиял на развитие его интеллекта.

По тем или иным причинам у этого животного развилась выдающаяся нервная система. Осьминоги могут использовать инструменты, решать задачи и демонстрируют неожиданные творческие подходы[16]. Классическим стал пример, в котором эти моллюски научились откручивать крышки стеклянных банок, чтобы добраться до лакомства внутри. У осьминога есть центральный мозг, а также небольшие независимые процессоры в каждом щупальце; таким образом получается уникальная комбинация централизованного и распределенного управления[17]. Также у животного, вероятно, есть модели самого себя: богатые, постоянно обновляющиеся сгустки информации для отслеживания своего тела и поведения. С инженерной точки зрения, чтобы функционировать эффективно, ему бы пригодились эти модели. Например, у моллюска может быть что-то вроде схемы тела, которая следит за его формой и структурой, чтобы координировать движения (возможно, у каждого щупальца есть своя схема себя). В этом смысле можно сказать, что осьминог знает о самом себе. Он обладает как этой информацией, так и сведениями об окружающем мире, и эти данные приводят к сложному поведению.

Но перечисленные действительно чудесные черты не означают, что у осьминога есть сознание.

Исследователи сознания иногда используют термин “объективное осознание” для обозначения того, что информация попала внутрь, обрабатывается и может повлиять на выбор поведения[18]. Это определение задает невысокую планку: так можно сказать, что микроволновая печь осознает настройки времени, а беспилотный автомобиль — надвигающееся препятствие. Да, осьминог объективно осознает себя и объекты вокруг. В нем содержится информация.

Но осознает ли он *субъективно*? Если бы осьминог умел говорить, мог бы он сообщить о субъективном опыте сознания так же, как мы с вами?

Давайте его и спросим. Проведите неправдоподобный мысленный эксперимент (и запомните его — он нам еще пригодится в этой книге). Предположим, в нашем распоряжении оказался потрясающий научно-фантастический прибор — назовем его Речинатор-5000, — который переводит информацию в речь. В нем есть порт, к которому можно подключить голову осьминога, и прибор вербализует информацию, найденную в мозге.

Прибор может озвучить что-то вроде: “Там рыба”, если зрительная система осьминога содержит информацию о рыбе, плывущей неподалеку. Он может сказать: “Я существо с кучей конечностей, которые могут двигаться так и сяк”. Или: “Чтобы достать рыбу из банки, нужно повернуть ту круглую штуку”. Прибор бы многое сказал, отражая информацию, которая, как мы знаем, содержится в нервной системе осьминога. Но нам неведомо, произнесет ли он: “У меня есть субъективный личный опыт — осознание — этой рыбы. Я не просто обрабатываю информацию о ней. Я ее *переживаю*. Я *чувствую*, каково это — видеть

рыбу”. Мы не знаем, есть ли в мозге информация подобного рода, поскольку не в курсе того, что сообщают осьминогу его модели самого себя. У него, возможно, нет механизмов, чтобы смоделировать сознание или приписать себе это свойство. Применение понятия “сознание” по отношению к этому животному может оказаться нерелевантным.

Тайна осьминога — пример того, что животное может быть сложным и умным, а мы тем не менее все еще не в силах ответить на вопрос о его субъективном опыте или даже о том, есть ли смысл задавать такой вопрос применительно к этому существу.

Возможно, один из источников путаницы здесь — невольное, но мощное стремление человека приписывать сознание всему вокруг. Как я подчеркнул в первой главе, мы склонны видеть сознание у кукол и других, еще менее вероятных кандидатов. Люди иногда верят, что их домашние растения осознают. Осьминог, у которого богатый поведенческий арсенал и большие глаза, наполненные сфокусированным вниманием, является в некотором роде тестом Роршаха с чернильными пятнами, убедительно запускающим в нас сильное социальное восприятие. Мы не только умом понимаем, что он собирает объективную информацию о мире, — мы не можем не чувствовать, что из этих задумчивых глаз исходит субъективное осознание. Но правда состоит в том, что мы этого не знаем, и наше ощущение сознающего разума говорит больше о *нас*, чем об осьминогах. Специалисты, которые изучают осьминогов, рискуют стать самыми ненадежными экспертами, потому что именно на них прежде всех остальных подействуют чары этих удивительных созданий. Позже, в пятой главе, я вернусь к всепроникающему аспекту человеческого сознания: оно инструмент в нашем

социальном арсенале, и мы безотчетно приписываем его тем, кто действует вокруг нас.

Чтобы внести ясность: я не утверждаю, что у осьминогов *нет* сознания. Но нервная система этих моллюсков до сих пор настолько неполно изучена, что мы не можем сравнить организацию их мозга с организацией нашего и предположить, до какой степени могут быть похожи на наши их алгоритмы и модели самих себя. Для проведения подобных сравнений нам нужно заняться животными из своей собственной родословной — позвоночными.

Глава

3

Централизованный интеллект лягушки

В детстве я много времени проводил на ферме на севере штата Нью-Йорк. Каждое лето целыми ночами мы слушали брачное кваканье лягушки-быка в пруду за домом. Мы звали его Элвисом, а лягушку, чей голосок потоньше доносился в ответ, — Присциллой. С тех пор я обожаю лягушек, а занявшись нейробиологией, захотел узнать, что происходит у них в головах.

У этих животных есть область мозга, которая называется “тектум”. На латыни это значит “крыша”, тектум — крыша среднего мозга, самый заметный выступ на его верхушке. Он есть не только у лягушек. Возможно, лучше всего он изучен у амфибий, но присутствует также у рыб, рептилий, птиц и млекопитающих. Эта область мозга есть у всех позвоночных, и, насколько нам известно, ни у кого другого. Можно с немалой уверенностью предположить, что тектум развился примерно полмиллиарда лет назад у маленьких бесчелюстных рыб, общих предков позвоночных, и все потомки унаследовали эту часть мозга[1].

У людей тоже есть тектум, но у нас он расположен не на верхушке мозга. Это сравнительно небольшой выступ (точнее, их два — по одному с каждой стороны), погребенный под кипами мозговых структур, которые расширились в нашем эволюционном прошлом. У людей и других млекопитающих он обычно называется верхним холмиком четверохолмия. Здесь для простоты я буду называть этот холмик тектумом.

Бóльшую часть эволюционной истории позвоночных тектум был вершиной интеллектуальных достижений: самый сложный процессор в центре мозга. У лягушки он принимает зрительную информацию и выстраивает из мира вокруг амфибии некий аналог карты[2]. Каждая точка на округлой поверхности тектума соответствует точке в окружающем животное пространстве. Тектум с правой стороны мозга лягушки содержит точную карту зрительного поля левого глаза, то же самое с левым тектумом и правым глазом. Когда вокруг лягушки хаотично летает черная точка, глаза принимают эту информацию, зрительный нерв посылает сигналы в тектум, а тот запускает управление мышцами. В результате язык лягушки “выстреливает” с потрясающей точностью и ловит муху.

Логику такого устройства ввода-вывода особенно ярко продемонстрировал нейробиолог Роджер Сперри. В начале 1960-х гг.[2] он провел на лягушке операцию: отделил глаза, перевернул их на 180° и вставил обратно[3]. Глаза прижились. У лягушек удивительные способности к регенерации. Зрительный нерв заново пророс от глаз к тектуму и восстановил внутреннюю зрительную карту. Когда

подопытная лягушка вновь начала видеть, при появлении мухи над головой она стала выбрасывать язык вниз. Если муха жужжала справа от лягушки, язык вылетал влево. Централизованный интеллект лягушки — это простой, но идеально эффективный механизм, который собирает сигналы от нервов и подбирает для них соответствующие реакции. К сожалению, манипуляции ученых его обманули. Модифицированную лягушку пришлось кормить с рук, иначе она бы погибла от голода.

Тектум лягушки занят не только зрением. Он также собирает информацию от ушей и осязательных рецепторов на коже[4]. Карта поверхности тела лягушки, а также слухового и зрительного пространств вокруг животного сходятся и частично интегрируются в тектуме. Это высший уровень интеграции в мозге амфибий: центральный процессор, который собирает воедино разрозненные сигналы, поступающие из окружающей среды, сосредоточивается на самом важном событии, происходящем в каждый конкретный момент, и запускает реакцию[5]. Тектум — механизм централизованного внимания лягушки.

Ученые могут прощупывать мозг с удивительной точностью, подобно тому как инженер-компьютерщик прощупывает микросхему. В одном из стандартных методов используются электроды: тонкие, как волосок, жесткие проводки, покрытые пластиковой изоляцией везде, кроме кончика. Оголенной остается примерно десятая доля миллиметра провода. Словно миниатюрный детектор, электрод в состоянии обнаруживать электрическую активность на микроскопическом расстоянии от оголенного металла. Длинный, гибкий провод, тянущийся от

электрода, соединяет его с принимающим оборудованием. Точный механизм закрепляет электрод на месте, а затем двигает его микрометр за микрометром, чтобы исследовать заданную область мозга.

Такая схема достаточно чувствительна для измерения активности отдельных нейронов в мозге. Когда нейрон вблизи кончика электрода подает сигнал своим соседям, устройство регистрирует этот крошечный электрический импульс. Сигнал усиливается и передается в динамики, а экспериментатор слышит щелчок. В обычных обстоятельствах нейрон выдает один-два случайных щелчка в секунду, но, если он активно задействуется в происходящем, клетка может внезапно разразиться сотней щелчков за секунду. Любимая забава нейробиологов — слушать щелчки отдельных нейронов и гадать, какую роль те выполняют в мозге.

Каждый нейрон в тектуме лягушки работает как детектор[6]. Он следит за определенной зоной пространства — например, областью непосредственно над головой — и срабатывает чаще, когда в эту область попадает какой-то объект. Нейроны бывают разные: какие-то предпочитают движущиеся определенным образом зрительные стимулы, другим больше нравятся звуки или прикосновения. По крайней мере некоторые нейроны мультисенсорны: для них нет разницы, приближается к макушке видимый объект, раздается оттуда звук или к голове прикасаются, — они сработают, чтобы передать сигнал остальному мозгу. Если два или более чувств сходятся, передавая одно и то же сообщение о приближающемся объекте, соответствующие нейроны в тектуме становятся особенно активными. Простое

вычисление словно говорит: “одна улика — уже хорошо, а если их две или три — явно происходит что-то важное”[7].

Подобный экспериментальный метод можно использовать и в обратном направлении: посылать по электроду импульсы, чтобы активировать близлежащие нейроны. Этот метод называется микростимуляцией. Такая стимуляция настолько слаба, что на коже вы ее не почувствуете, но ее хватает, чтобы пощекотать нейроны и побудить их послать свои собственные сигналы. Использование микростимуляции позволяет задать вопрос: “Если искусственно заставить возбуждаться эту группку нейронов у кончика электрода, что они велят делать животному?”

Скажем, саламандра при электрической стимуляции тектума производит сложное скоординированное движение[8]. Она поворачивается, открывает рот, высовывает язык, вытягивает передние конечности и делает хватательные движения своими длинными тонкими пальцами — будто ловя добычу. Какую бы область пространства ни отслеживали нейроны в определенной зоне тектума, при электрической стимуляции этих нейронов животное будет тянуться к той самой области.

Стимулируйте точку на карте тектума игуаны — и повернутся ее тело, голова, глаза[9]. Животное будет смотреть ровно на то место, которому соответствует ваша точка на карте.

Стимулируйте тектум рыбы, и ее тело изменит положение, чтобы сориентироваться на нужную область пространства[10]. Точно развернуться в нужном направлении для рыбы — это не просто пошевелить

шейным суставом. Здесь требуется сложное взаимодействие плавников и воды.

У гремучих змей есть своя версия инфракрасного зрения: пара специализированных чувствительных к температуре органов, расположенных посередине между глазами и ноздрями. Эти органы посылают информацию в тектум, который формирует карту температурных сигналов, наложенную на обычную зрительную карту пространства[11]. Предполагается, что на этой мультисенсорной карте основываются как способность змеи поворачивать голову в сторону добычи, так и точность ее нападения.

В тектуме совы зрительная карта совмещена со звуковой[12]. Когда птица охотится, она может нацеливаться, либо увидев добычу, либо, при охоте ночью, услышав ее шуршание в траве.

Стимулируйте верхний холмик обезьяны, и произойдет стремительное скоординированное движение головы и глаз[13]. Обезьяна повернется к нужной точке пространства. Мне не встречались исследования с применением электрической стимуляции к верхнему холмику мозга человека, но мы — подвид приматов, и у нас предположительно действует тот же механизм, что и у обезьян. Когда вы поворачиваетесь на что-то посмотреть, особенно если неожиданное событие заставляет вас ориентироваться быстро, рефлекторно, — это непринужденное на вид, хорошо скоординированное движение скорее всего запускается из тектума.

Все позвоночные пользуются тектумом примерно одинаковым образом, хотя у многих видов есть свои дополнительные особенности. Область мозга собирает сенсорную информацию, выбирает самое яркое из происходящего вокруг и направляет животное,

физически поворачивая его органы чувств в нужную сторону.

Такая ориентировка иногда называется явным вниманием[14]. Это простое решение фундаментальной проблемы: вокруг происходит так много всего, что мозгу не справиться с обработкой всей информации. Животному нужно выбрать наиболее его интересующее и отбросить остальное. Если вы направите глаза и уши на один объект, то автоматически отбросите другие события, которые окажутся на периферии. Для вас эту работу выполняет тектум. Это первый в эволюции “центральный пульт управления” вниманием в мозге позвоночных.

Большинство людей, говоря о внимании, имеют в виду именно явное. В обиходном смысле слова, на что вы смотрите — тому и уделяете внимание. Отвернувшись от объекта — не уделяете.

Но взгляд — это лишь часть истории о внимании. Студент может машинально черкать на бумажке, смотреть в тетрадь, но по-прежнему обращать *скрытое* внимание на преподавателя. Или представьте, что вы случайно услышали, как люди вас обсуждают. Вы не станете поворачиваться к ним, чтобы не выдать себя, но ваше внимание, ваши ресурсы обработки информации сосредоточатся на этом разговоре. Или вы можете замечтаться, сидя в кресле, и ваше внимание обратится на что-то, чего попросту не существует в физическом мире, а ваш взор будет рассеянно блуждать по потолку. Во всех этих примерах направление внимания не совпадает с направлением взгляда. Этот более сложный его вид — скрытое внимание — не входит в обязанности тектума, который занимается только явной ориентировкой. С тектумом в роли основного центра

внимания лягушка в состоянии пользоваться только явным вниманием. Она может физически разворачиваться к объектам окружающего мира.

Во внимании — явном ли, скрытом — нет смысла, если им нельзя управлять. Но управление — не такая уж простая инженерная задача. Нужно тщательно отслеживать управляемое. Впервые в этой эволюционной истории мы встретим не просто клетки, обрабатывающие информацию, и не просто животных, направляющих внимание, но мозговые системы, которые создают *схему внимания* — комплекс информации (его называют внутренней моделью), следящий за состоянием внимания. Наша эволюционная история подбирается все ближе к чему-то напоминающему сознание. Но пока еще не добралась.

Беспилотному автомобилю нужна внутренняя модель всей конструкции. Встроенный в него компьютер должен не только получать информацию о внешнем мире и затем посылать сигналы рулю и педалям. Системе необходима информация о самой машине, ее форме и размере, ее поведении на дороге, ее постоянно меняющихся характеристиках: скорости, ускорении, местоположении. Без богатой, постоянно обновляемой внутренней модели, содержащей большой объем информации, у машины будет лишь центр управления, который посылает водительские команды, но, скорее всего, дело кончится аварией.

Принцип внутренней модели был впервые описан в инженерной сфере[15]. Неважно, что управляется — что-то материальное, вроде машины или роботизированной руки, или нечто аморфное, например поток воздуха во всех помещениях большого здания. Чтобы система управления работала

как следует, ей нужна внутренняя модель того, чем она управляет. Ей требуется возможность наблюдать машину, робота или потоки воздуха. Внутренняя модель чем-то напоминает карту на столе генерала — с маленькими пластиковыми танками и солдатиками. Это связный комплекс информации, который, обычно упрощенным или схематичным образом, отражает и отслеживает то, чем нужно управлять.

Тот же принцип работает и в биологии. Мозг управляет телом с помощью внутренней модели, так называемой схемы тела — комплекса информации о его структуре и постоянно меняющемся состоянии[16]. Иногда при инсульте повреждаются области мозга, которые строят схему тела[17]. Если пациент больше не осознает форму или структуру своей руки, он не сможет ею управлять. Пострадают простые навыки — указывать на что-то, дотягиваться рукой, держать чашку. Но увидеть важность внутренней модели можно и не заглядывая в отделение постинсультной реабилитации. Повесьте тяжелую сумку с покупками на запястье и попробуйте взяться за ручку двери: поначалу ваши движения будут неуклюжими. Внутренняя модель руки, имеющаяся у мозга, внезапно оказывается неправильной: изменились динамические свойства конечности. Но очень быстро, за несколько попыток, внутренняя модель выучит новые правила, и ваши движения станут плавнее и точнее[18].

С инженерной точки зрения внутренняя модель должна отслеживать настоящее *и* предсказывать будущее. Если вы хотите чем-то управлять, например тележкой в магазине, нужна возможность предсказать, что она будет делать в следующую секунду. Вы создаете что-то вроде интуитивного

симулятора тележки, запускаете его в голове и понимаете, как она себя поведет. То, как вы станете управлять реальной тележкой, какую силу и под каким углом приложите к ее ручке, будет зависеть от предсказаний, сделанных внутренней моделью. Дети плохо справляются с подобной задачей и врезаются в магазинные полки: отчасти это происходит потому, что у них не сложилась хорошая внутренняя модель тележки. Они не могут предсказать, как усилие, приложенное к ручке, повлияет на движение колес. Взрослые же, попрактиковавшись, вырабатывают бессознательную, интуитивную модель.

А как обстоят дела со вниманием? Это ведь, можно сказать, важнейший процесс в мозге, и, несомненно, им нужно управлять. Чтобы эффективно реагировать на мир, мозг должен уметь стратегически сосредоточивать ресурсы на произвольных предметах. Но при этом внимание бывает капризным и расхлябанным не менее, чем тележка в супермаркете, сбиваясь куда попало. Из базовых принципов инженерии управления мы знаем, что тектуму нужна внутренняя модель, чтобы следить за вниманием. Мы с коллегами назвали эту предположительную внутреннюю модель "схема внимания" — по аналогии со схемой тела, которая помогает следить за телом. Схема внимания — это комплекс информации, описывающий внимание: не предмет, на который оно направлено, а само внимание. Схема наблюдает за его состоянием, отслеживает его динамические переходы от одного состояния к другому и предсказывает, как оно может измениться в ближайшем будущем. Вариант схемы внимания — информация, которая следит конкретно за явным вниманием, — был обнаружен в тектумах обезьян и кошек[19]. Согласно базовым принципам, такая же информация почти

наверняка имеется у лягушек, рыб и других животных, у которых есть тектум, даже если он подробно не изучен.

Вернемся к нашим лягушкам. Мы знаем, что у них есть центральный процессор — тектум. Мы знаем, что им присуще явное внимание — способность ориентировать органы чувств на определенный фрагмент большого мира. Мы знаем, что у них должна быть схема внимания, поскольку во внимании нет смысла, если им невозможно управлять, а управлять им невозможно без внутренней модели. Схема внимания — это примерно то же самое, что сложная модель самого себя. Лягушка не просто направляет внимание на определенные предметы в своем мире. Она еще и некоторым образом знает, что делает это. У нее есть информация *о* ее собственном внимании.

Что именно знает о себе мозг лягушки благодаря схеме внимания?

Вспомним о мысленном эксперименте, который я предложил провести во второй главе. Представим себе, что мы взяли футурологический переводчик информации в речь, Речинатор-5000, и воткнули его в тектум лягушки. Пользуясь информацией из схемы внимания, Речинатор может сказать: “Тут какие-то глаза. Тут тело. Они двигаются так и этак, поворачиваясь в разных направлениях. В данный момент они направлены на ту дергающуюся черную точку. Поскольку прямо сейчас они двигаются, то скоро окажутся направлены вон в ту сторону”. Информация настолько буквальна потому, что внимание лягушки ограниченно. Да, у нее есть схема внимания, но она описывает лишь явное. Для лягушки внимание — это поворот головы и глаз. А значит, нужная лягушке внутренняя модель — это модель

головы и глаз: как они двигаются, как соотносятся с предметами.

Предположим, мы при помощи Речинатора спросили у тектума лягушки: “А есть ли у тебя субъективное *переживание* этой мухи?”

Тектум может давать только ту информацию, которая в нем имеется. Он скажет: “Там носится туда-сюда черная точка. Тут глаза. Тут тело. Они двигаются. Они направлены туда”.

Мы слегка раздражаемся: “Да-да, это понятно. Но как насчет *осознания*? Насчет *мысленного образа* мухи?”

Тектум лягушки повторяет: “Тут глаза. Тут тело. Они направлены туда”.

В тектуме лягушки попросту нет информации, чтобы ответить на наши вопросы. Понятие осознания для нашей амфибии не имеет смысла. Несмотря на то что у нее есть сложный мозг, определенный тип внимания и даже схема внимания, лягушке не нужны внутренние модели, описывающие ее как сознающего агента.

Я до сих пор с теплом вспоминаю Элвиса и Присциллу. Я знаю, что их поведение на удивление сложно, даже их скрипучие брачные песни. Если бы я достаточно времени провел с лягушками, у меня бы наверняка наладился с ними контакт и появилось бы интуитивное ощущение (столь характерное для нас, общественных человеческих существ), что в этих крошечных животных должно таиться сознание. Это человеческие, социальные объяснения того, почему людям может казаться, что у лягушки есть сознание. Но у амфибии практически наверняка отсутствует аппарат, позволяющий моделировать сознание или

приписывать это свойство себе самой. У нее может быть *объективное* осознание о себе и окружающей среде в том смысле, что она обрабатывает информацию о своем теле и его окружении, но, если бы нам удалось перевести эту внутреннюю информацию в речь, не нашлось бы никаких причин обнаружить в лягушке *субъективное* осознание.

И тем не менее все нужные фрагменты почти расположились по местам. Согласно моему эволюционному отчету на настоящий момент, полмиллиарда лет назад у древней бесчелюстной рыбы образовалась некая форма явного внимания, появился тектум, чтобы этим вниманием управлять, и, вероятно, схема внимания, чтобы упростить управление. Амфибии, рептилии, птицы и млекопитающие — все унаследовали одну и ту же систему. Во всех нас скрыт один и тот же тектальный аппарат. Но, чтобы обнаружить феномен, который мы признаем как сознание, нужно сделать еще один шаг. Нужно от явного внимания обратиться к более сложному и тонкому навыку скрытого, в котором экспертами являются птицы и млекопитающие.

Глава

4

Кора головного мозга и сознание

Кора головного мозга — характерная для мозга млекопитающих складчатая внешняя оболочка — начала развиваться, вероятно, у рептилий более 300 млн лет назад, в каменноугольном периоде (карбоне)[1]. То были времена бескрайних топких джунглей, мир все еще объединялся в суперконтинент Пангею[2], кислорода в воздухе содержалось больше, чем сегодня, и от этого насекомые вырастали до гигантских размеров, несмотря на невысокую эффективность их дыхательных систем[3]. С тех времен в окаменелостях сохранились двухметровые многоножки и полуметровые стрекозы. Период называется каменноугольным, поскольку после увядания пышных джунглей весь огромный объем биомассы превратился в угольные пласты.

Ранний каменноугольный мир, скорее всего, кишел беспозвоночными и амфибиями — единственными на тот момент животными, которые приспособились к жизни на суше[4]. Амфибии эволюционировали в рептилий к концу каменноугольного периода, когда климат стал менее влажным[5]. У них появился водонепроницаемый

чешуйчатый покров, а у яиц рептилий — чтобы можно было откладывать их вдали от воды — твердая скорлупа.

Также у них развилась часть мозга, которую называют Wulst (гиперпаллиум конечного мозга)[3]: это вырост его наиболее фронтальной части, именуемой передним мозгом[6]. Wulst у рептилий выполняет главным образом сенсорную функцию. Он содержит специфическим образом организованную карту зрительного пространства и карту осязательных рецепторов на коже — наподобие сенсорных карт в тектуме, о которых я писал в предыдущей главе. Wulst — что-то вроде тектума 2.0, радикальная эволюционная инновация; это как айфон в сравнении со старым дисковым телефоном.

Рептилии незаслуженно “прославились” своей глупостью. “Рептильный мозг” в расхожем представлении — это древнее ядро, таящееся за разумными частями. Но на самом деле у этих животных уже есть зачатки того, что станет корой головного мозга у человека. В них и правда есть искра интеллекта. Многие виды рептилий умеют решать сложные задачи и вступать в сложные социальные взаимодействия[7].

Вскоре после своего появления в позднем карбоне рептилии разделились на две основные группы. Синапсиды — так называемые звероподобные — вначале несильно отличались от представителей другой группы, завропсид[8]. У синапсид была немного иной структура черепа, которая обеспечивала более надежную опору мышцам челюсти. Проще говоря, у них лучше получалось питаться, а это давало им энергетическое преимущество. Самый известный из ранних синапсид — диметродон пермского периода.

Это доисторическое создание, фигурка которого есть в наборе динозавров каждого ребенка: посадка низкая, как у крокодила, на спине большой ребристый парус. Но это не динозавр. Это звероподобная рептилия, которая ближе к нам, чем к тираннозавру. Вероятно, она была выше и меньше стелилась пузом по земле, чем может показаться при взгляде на убогие пластиковые игрушки.

Шло время — сотни миллионов лет, звероподобные рептилии постепенно превращались в современных млекопитающих, а Wulst разрастался в многослойную оболочку из нейронов над остальным мозгом[9]. У многих млекопитающих кора — это гладкое наружное покрытие мозга. Но у некоторых видов кора так разрослась и площадь ее поверхности настолько увеличилась, что ей пришлось собраться в складки — кора стала напоминать утрамбованный в черепе ком ткани. Разглаженная, она размером и толщиной походила бы на большое махровое полотенце.

Другие связанные с корой структуры в мозге также разрослись: особенно это касается крупного, напоминающего авокадо по форме участка в основании мозга, который называется “таламус”[10]. Название ему подходит: *thalamus* на латыни значит “ложе” или “основание” (изначально — “спальня” по-гречески). С таламусом соединяется каждая область коры, и почти вся поступающая в нее информация сначала проходит как раз через таламус, отчего его прозвали вратами в кору. Ведя речь об этой специализированной сети, для точности следовало бы пользоваться неуклюжим техническим термином “таламо-кортикальная система”, но для простоты я ограничусь сокращенным “кора”.

Пока у звероподобных рептилий в течение 300 млн лет развивалась сложная кора головного мозга, их незвероподобные собратья (зауропсиды) нашли другой путь к повышению интеллекта. Этот путь вел от животных, которые называются архозаврами, к динозаврам и, наконец, современным птицам. Архозавры были крупными хищными рептилиями с низкой посадкой и чуть более крупным мозгом, нежели у их непосредственных предшественников[11]. Крокодилы — современные архозавры — одни из умнейших и поведенчески сложно организованных рептилий[12]. Об их интеллекте наглядно свидетельствуют примеры того, как они искусно подстерегают добычу, делятся едой, воспитывают детенышей.

Примерно к 230 млн лет назад, в триасовом периоде, архозавры образовали специфическую подгруппу причудливых созданий, которых эволюция заставила бегать на удлиненных задних ногах[13]. К удивлению тех, у кого в голове живут образы гигантских четвероногих динозавров, ископаемые находки показали, что все динозавры произошли от единого предка, который ходил на двух ногах и, возможно, иногда касался земли передними конечностями. В ходе дальнейшей эволюции некоторые травоядные виды вновь опустились на четвереньки, а хищные динозавры остались прямоходящими.

Всем знаком стереотип, что мозг динозавра размером с орех. Это самая что ни на есть клевета. Внутренние слепки ископаемых черепов кое-что сообщают о размере и структуре его мозга[14]. У самых крупных динозавров и мозг был соответствующих массы и размера. К примеру, у тираннозавра этот

орган мог по размеру соперничать с человеческим, хотя в нем, скорее всего, было намного меньше нейронов и связей между ними[15]. Мозг динозавров, видимо, был устроен примерно как у крокодилов, включая Wulst — ту самую протокору. Они не были глупыми. Особенно тероподы, двуногие хищные динозавры, — на тот момент они, пожалуй, были умнейшими животными на планете. Их обращенные вперед глаза давали стереоскопическую глубину зрения: вероятно, для лучшей обработки потока входящей информации у них разрослась зрительная часть области Wulst.

Птицы, как известно каждому современному школьнику, произошли от динозавров. Но этот популярный научный факт — упрощение более тонкой и странной правды. Птичий полет возник и постепенно развился, видимо, в юрском периоде, между 200 и 145 млн лет назад[16]. Пожалуй, не имеет смысла указывать конкретную точку во времени и говорить: “Вот тогда-то и появились птицы”. Нечеткость грани между динозаврами и птицами особенно заметна при рассматривании найденного в китайской провинции Ляонин великолепно сохранившегося в вулканическом пепле ископаемого пернатого динозавра[17]. Сейчас складывается представление, что птицы *и есть* динозавры-тероподы. Живи вы в мезозойскую эру, вам бы никогда и в голову не пришло выделять птиц как отдельную категорию. Мир был полон пернатых динозавров: крупных и мелких, зубастых и клювастых, прогуливающихся на двух ногах или пролетающих над вашей головой. Если бы мимо вас пробежала современная курица или рядом вспорхнул воробей, вы бы подумали: “Смотри-ка, еще один вид динозавров-теропод”. Но вы бы никогда не сказали: “Глядите,

динозавры породили новый класс животных!” Так же как сегодня никто бы не сказал, что летучие мыши произошли *от* млекопитающих. Летучие мыши *и есть* млекопитающие, которые приспособились к полету. А птицы, получается, и есть летучие динозавры, дожившие до наших дней. И нейробиологи весьма активно их изучают.

Wulst у птиц значительно расширен по сравнению с аналогичным органом рептилий[18]. За сотни миллионов лет, в течение которых ветвь млекопитающих колдовала над своим мозгом, выращивая из рептильного Wulst тонко проработанную кору, ветвь динозавров/птиц тоже не дремала, находя свои способы расширить и улучшить ту же самую, изначальную структуру мозга.

И хотя выражение “птичьи мозги” все еще звучит оскорбительно, птицы — тоже умные существа. У некоторых видов есть сложноорганизованная социальная жизнь, некоторые — искусные охотники, некоторым присуща блестящая память на места хранения еды. В особенности славятся своим умом вороны. Они могут гнуть из проволоки крючья, чтобы подцепить немножко пищи[19]. Сюжет знаменитой басни Эзопа о вороне, которая бросала камни в кувшин, пока уровень воды не поднялся так, что она смогла достать плавающий там кусочек еды[4], — не выдумка. Вороны и вправду справляются[5] с этой задачей[20].

Похоже, что своим феноменальным умом птицы обязаны разросшемуся переднему мозгу. Вряд ли интеллект вымерших динозавров шел хоть в какое-то сравнение с нынешним птичьим: современные вороны представляют собой недавнее достижение эволюции, но у всей линии рептилии/динозавры/

птицы прослеживаются одни и те же базовые мозговые структуры.

Далее я сфокусирую внимание в основном на коре головного мозга млекопитающих, поскольку мы намного больше о ней знаем и поскольку млекопитающие — единственные животные, которые на сегодняшний день могут словесно засвидетельствовать наличие сознания (это мы, люди). Но неплохо бы иметь в виду, что другие виды животных — птицы, крокодилы — располагают сходными мозговыми системами, благодаря которым возникает сложное и разнообразное поведение.

Теперь, когда я сжато описал эволюцию коры головного мозга, я объясню, почему считаю, что сознание основывается именно на этой структуре. Для этого мне нужно будет привести некоторые подробности обработки информации в коре. Это не просто увеличенный процессор вроде тектума из мозга лягушки, прилепленный поверх меньших, более ранних моделей. Это принципиально иной *вид* процессора — машина, предназначенная, чтобы просеивать огромные объемы информации и отбирать оттуда лишь небольшое ее количество. Оно, в свою очередь, проходит тщательную глубинную обработку и в конечном итоге направляет поведение. В силу этого непрерывного отсеивания кора предстает, в сущности, машиной внимания.

Кора головного мозга чем-то напоминает Национальную футбольную лигу (НФЛ). Команды в НФЛ проходят ряд матчей-соревнований, пока Супербоул[6] не определит окончательного победителя. В коре информация фильтруется по иерархическим уровням, в постоянном соревновании по принципу “победитель получает всё”[21]. Независимо

от модальности информации — зрительной, слуховой, эмоциональной, интеллектуальной — архитектура коры создает соревновательные раунды на выбывание. Отбираемые со все возрастающей придирчивостью сигналы проходят более глубокую обработку и получают больше шансов повлиять на поведение. Наконец, некий комплекс информации побеждает в корковом Супербоуле, и (хотя бы на миг) вся обработка сосредоточивается на нем.

Такой способ обработки-на-выживание особенно хорошо изучен для зрения. На рис. 4.1 показана частичная схема зрительной системы, которую удалось построить благодаря многим десятилетиям исследований на приматах — как обезьянах, так и людях[22]. Свет входит в глаз и образует картинку на сетчатке — светочувствительном слое ткани в глубине глаза. Нейроны сетчатки активируются и немедленно начинают соревнование местного уровня. Сигналы соперничают с соседями, более сильные подавляют более слабых. Особое преимущество получают те, что соответствуют резким зрительным контрастам.

Затем информация покидает глаз по зрительному нерву (кабелю из примерно 1,5 млн волокон) и попадает в таламус у основания мозга. Входящая зрительная информация направляется в определенную часть таламуса — выступ на его боку, который называется “латеральное коленчатое тело”. Там она снова процеживается сквозь соревновательное сито.

Пройдя латеральное коленчатое тело, информация отправляется по волокнам на первый уровень зрительной коры, в область затылочной доли мозга, которая называется первичной зрительной корой, или

V1, где снова происходит вышеописанное соревнование нейронов.

Из V1 данные идут в хитросплетение высших зрительных областей — V2, V3, V4, MT, MST, TEO, TE и т.д. — в буквенный суп вместо названий. За последние полвека нейробиологи разметили зрительную кору во все более мелких подробностях, выделив десятки зон, в общей сложности покрывающих 60т.о. коры, в основном в затылочной, височной и теменной долях. В каждой из этих зон размером с почтовую марку фрагменты информации постоянно соревнуются — нейроны тормозят своих соседей.

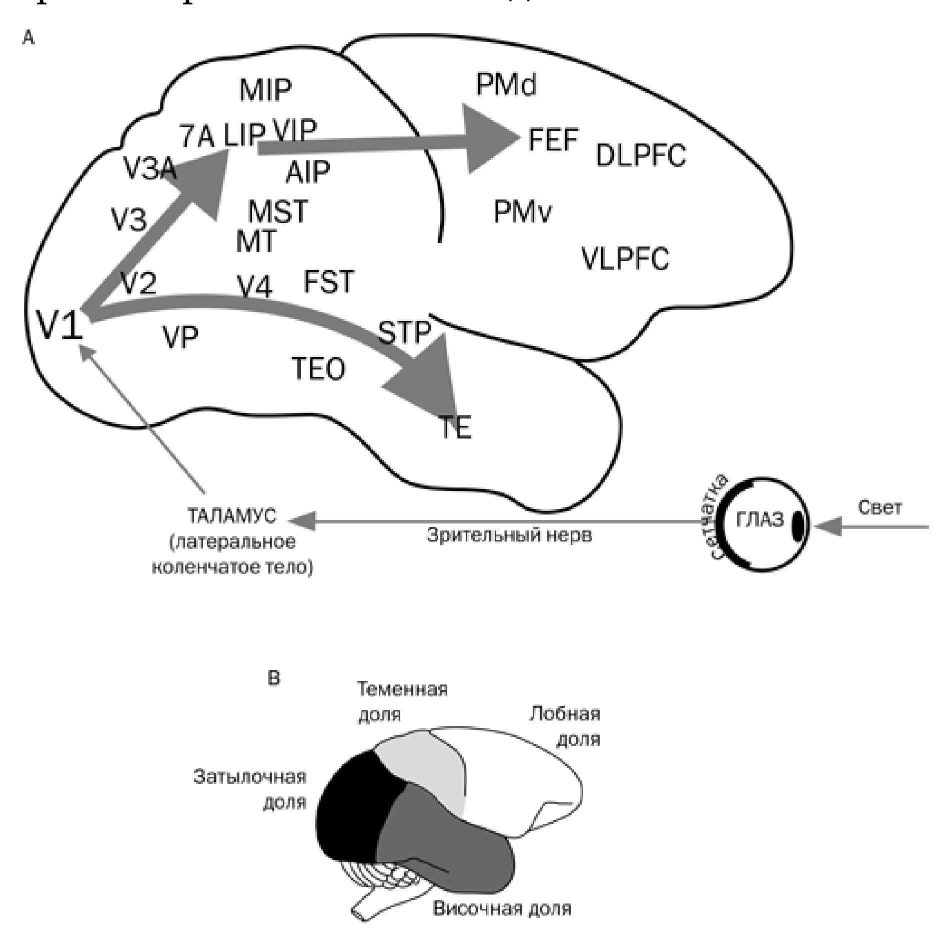


Рис. 4.1. Некоторые зрительные зоны коры в мозге примата. (A) Основные области, задействуемые в зрении и прохождении информации через эти зоны в целом. (B) Четыре доли головного мозга

Зрительные области коры в целом организованы так, что сложность повышается по мере того, как

данные текут от низкоуровневых зон в задней части мозга к зонам высоких уровней ближе к передней. К примеру, нейроны в V1, внизу иерархии, раскладывают зримый мир на контуры и цветовые пятна. Там происходит простое, поверхностное, но подробное разбиение видимого мира. Напротив, нейроны в TE, зоне намного выше в иерархии, чувствительны к сложным зрительным образам, таким как лица и руки. Они обрабатывают информацию о сущности объектов, а не о подробностях. И тем не менее данные текут не только вверх по иерархической лестнице. Сигналы могут идти обратно или вбок — в каких угодно направлениях по сети. Но и это еще не всё — вдобавок к сказанному выше кора также постоянно сообщается с глубинными структурами мозга, особенно с таламусом.

По мере того как информация течет по этой системе обработки, соревнование между сигналами все больше напоминает нечестный вариант НФЛ. Сильнейший входной сигнал побеждает не всегда. Другие источники могут обманом склонить чашу весов в свою пользу. Если посмотреть по сторонам, можно заметить самый яркий, самый быстрый объект, от которого исходит самый большой сигнал, но механизмы мозговой обработки необязательно берутся именно за него. Сигнал из лобной доли может быть перенаправлен обратно в зрительную кору, и фокус обработки сместится на менее очевидный стимул, который и победит в соревновании[23]. Когда выигрывает самый перцептивно значимый стимул снаружи, это называется восходящим (bottom-up) вниманием; когда сигнал изнутри дает взятку судье и состязание становится неравным, это называется нисходящим (top-down) вниманием. Но на самом деле это два разных аспекта одного и того же процесса.

В описании этих корковых разборок участвовал нейробиолог Роберт Десимон, на редкость уместно назвавший их предвзятым соревнованием[24]. Мне это видится одной из основных организационных истин о коре. Локальное торможение нейронами друг друга, из которого возникает соревнование, — главенствующий механизм в ее работе. Неслучайно эпилепсия — это заболевание коры[7]. Эпилептический припадок происходит, когда отказывает локальное торможение[25]. Сигналы, которые обычно сдерживают друг друга, внезапно начинают распространяться, и неконтролируемая активность захлестывает всю кору без разбору. Болезнь показывает, насколько важно торможение в коре и насколько катастрофичны последствия его сбоев.

Во второй главе я описал исполняемый в глазу краба незамысловатый трюк, который называется латеральным торможением — когда соседние нейроны подавляют друг друга[26]. Результат латерального торможения — повышение резкости изображения. Яркие участки кажутся ярче, темные — темнее. Предвзятое соревнование — корковая версия латерального торможения, увеличенная в миллионы раз, раздутая от локальной конкуренции в глазу до бурного мира состязаний на выбывание в этой многоступенчатой иерархии.

Эта иерархия не ограничивается зрением. Просачиваясь сквозь взаимосвязанные сети областей мозга, информация зрительная сталкивается с другими ее видами. Осязание, слух, управление движениями, абстрактное мышление — обширные группы сигналов сходятся в нескольких областях, которые, возможно, занимают самое высокое

положение в иерархии, играя роль общих интеграционных сетей в мозге. Эти области в основном располагаются в двух долях: теменной и лобной, и поэтому их не мудрствуя лукаво назвали теменно-лобными сетями[8][27]. Дошедшая до них информация выиграла Супербоул. Она достигла того, что философ Дэниел Деннет назвал “славой в мозге”[28]. Корковые механизмы раундов на выбывание отобрали из окружающего мира всего несколько предметов, которые теперь влияют на работу руководящих сетей коры и обладают большими шансами на то, чтобы организовать поведение и сохраниться в памяти. Но то, что находится в фокусе сейчас, может покинуть его через минуту. Соревнование не прекращается, и новые фрагменты информации пробиваются по нейронным сетям на высшие уровни обработки. Корковая система не знает покоя, в этом суть ее работы. Состязание никогда не приводит к устойчивым результатам.

Описанный безостановочный процесс отсеивания в коре — самый мощный механизм нашего внимания. Это наиболее изощренный инструмент, которым располагает мозг для понимания мира. Иногда меня спрашивают, почему я построил теорию сознания вокруг, казалось бы, такой узкой темы, как внимание. Внимание — не узкая тема. В нем сущность того, как кора головного мозга постигает мир.

Тектум — та примитивная структура мозга возрастом в полмиллиарда лет, которую я описывал в предыдущей главе, — главенствует в явном внимании. Он может поворачивать ваши глаза и уши подобно спутниковым тарелкам, чтобы собрать побольше информации о конкретной части мира. Но эволюция коры головного мозга открыла новую возможность

обработки окружающей среды: скрытое внимание. Вы можете уделять внимание чему-то не находящемуся в центре поля зрения.

Я не имею в виду, что кора никак не связана с движениями глаз[29]. Она играет большую роль в координации глаз и головы. Но корковое внимание не обязано устремляться туда, куда направлен взгляд. Вы можете нервно рассматривать свои шнурки, а корковое внимание при этом будет пригвождено к начальнику, который стоит перед вами. У вас может сильно и неприятно зачесаться поясница, которая завладеет всем вашим вниманием, пока вы с улыбкой смотрите прямо на друга, не слыша ни единого его слова. Вы можете пялиться на станицу книги, а внимание сосредоточится на каком-то воспоминании или мысли, которые ну никак не связаны с направлением взгляда.

Разница между явным и скрытым вниманием такова: явное вбирает объект органами чувств, скрытое — мощной моделирующей машинерией коры. Вы можете “направить” эту машинерию на что угодно — на конкретный предмет, находящийся перед вами, на что-то, располагающееся в стороне, или на нечто внутреннее — мысль, переживание.

Ученые иногда сравнивают скрытое внимание с прожектором[30]. У него есть яркий сноп лучей в центре, направленный на один-два предмета, и размытая периферия, в зону которой могут попадать другие объекты. Из сотен предметов, окружающих вас в каждый конкретный момент, корковый прожектор освещает лишь немногие, а остальные остаются совсем без внимания, вы их даже не замечаете. Если представлять прожектор, который может озарять только реальное окружение, аналогия оказывается

слегка натянутой: скрытое внимание в состоянии перемещаться по измерениям, никак не связанным с пространством[31]. Можно смотреть на картину Мондриана и так сосредоточиться на цветах и оттенках, что совсем упустятся и не запомнятся формы, хотя цвета и формы были в одном и том же месте в пространстве. Такое внимание не очень похоже на прожектор. С учетом этих соображений аналогия с прожектором впала в немилость. Лично мне она нравится, но только если считать, что луч прожектора блуждает в пространстве абстрактных измерений, а не только в трех измерениях физического пространства.

Кора головного мозга создает внутренний прожектор, позволяющий нам исследовать безграничный многомерный ландшафт, на котором в фокус обработки информации могут попадать и конкретные предметы, и самые абстрактные идеи.

Глубоко обрабатывать информацию невозможно без смещения фокуса внимания. Если бы вам пришлось перелопачивать все доступные данные одновременно, вам понадобился бы мозг размером с планету. Эволюция нашла оптимальный способ, как при ограниченных мозговых мощностях обрабатывать мир с умом. Предложенное ею решение состоит в том, чтобы подвергать информацию настолько суровым состязательным испытаниям, что в каждый конкретный момент обрабатывалась бы лишь небольшая ее доля, а при этом в систему был бы встроен изощренный механизм, управляющий смещением и подстройкой фокуса обработки. Внимание — ключ, открывающий комплексное понимание мира.

Ученые, пишущие об эволюции сознания, подчеркивают постепенное увеличение сложности мозга. Интуитивно появляется соблазн считать, что от этого-то и возникает сознание[32]. Согласно такой точке зрения, где-то в ходе эволюции нервная система стала настолько сложной, что перешла некий порог, пробудилась и обрела субъектность. Будь это так, вопрос сознания свелся бы к поиску порога, — но здесь мы ступаем на скользкую дорожку.

Мы можем начать с допущения, что рыбы — тупые автоматы, у которых сознания быть не может. Однако вскоре эксперт по рыбам рассеет наше заблуждение, описав неожиданное разнообразие их поведения[33], и мы окажемся готовы заключить, что они все-таки могут обладать сознанием. Пчелы — всего лишь насекомые с крошечными мозгами, но моделирующая сложность и у них такова, что иногда по уровню выполняемых задач они могут посоперничать с домашними собаками[34]. Мой знакомый, который держит дома тарантула, утверждает, будто у этого животного настолько сложное поведение, такая яркая личность и темперамент, что оно явно обладает сознанием. Затем в дискуссию тихой сапой начнет проникать целый набор человеческих эмоциональных предубеждений. Раз тараканы осознают, может быть, не стоит их давить? А проведу-ка я (в целях собственного удобства) границу на шкале нервной сложности чуть повыше прежней и предположу, что теперь у тараканов сознания нет. И это оставит меня с некоторыми противоречиями, которые прокрадутся в силу простой человеческой иррациональности.

Недостатки подхода, опирающегося на сложность, заключаются в том, что чем дальше вы углубляетесь в тему, тем более ускользающей и произвольной

становится граница этой сложности, тем сильнее она зависит от людских капризов. У одноклеточных животных есть свои собственные непростые методы обработки информации при помощи взаимодействующих химических сигналов. Если вы специалист по амебам и проводите жизнь, выслеживая их в микроскоп, то, видя всю сложность их поведения, вы, возможно, начнете отстаивать наличие сознания у амеб. У растений есть свои системы электрохимической передачи данных об окружающем мире — так, может быть, и растения обладают сознанием? Геном в ядре клетки вкупе со всей тяжелой химической артиллерией, управляющей генетическим кодом, — информационно насыщенная вычислительная система. Так не исключено, что у ядра клетки тоже есть собственное сознание?

Но мы пойдем еще дальше! Камни, вода, отдельные электроны наполнены информацией, колебанием состояний, чувствительны к своему окружению. Почему бы и им не обладать сознанием?

Начав с предположения, что сознание естественным образом возникает из сложной обработки информации, трудно не впасть в панпсихизм — убеждение, что всё во Вселенной в какой-то степени обладает сознанием[35]. Заменяя сознание информацией и сложностью, которые можно найти буквально везде и во всем, мы оказываемся на нечеткой шкале, по которой можно скользить и вверх и вниз.

Я предполагаю, что сознание поддерживается корой головного мозга. Но не потому, что кора — самый сложный механизм в изученной Вселенной, хотя это так. Причина намного более конкретна.

Представьте себе, что вы смотрите на самый обычный предмет — скажем, яблоко. Оно лежит на столе перед вами, вам его хорошо видно. Зрительная информация о фрукте течет по вашей системе обработки данных от сетчатки в кору. Предположим, что сигналы от яблока побеждают в соревнованиях на выбывание. Взбираясь выше по корковой иерархии, информация о нем обрабатывается все глубже, а сигналы-конкуренты подавляются. И вот информация о яблоке добирается до центральных сетей в мозге — теменно-лобных. Десятки тысяч других возможных фрагментов информации в этот момент проиграли в соревновании. Звуки и видимые предметы вокруг вас, ощущение одежды на теле, воспоминания, мысли, эмоции — все это уступило сочному плоду. В данный конкретный момент времени корковая машина внимания захвачена зрительным образом яблока (и, возможно, нескольких других предметов), а информационные соперники обрабатываются лишь поверхностно, в гораздо более ограниченных областях вашей коры.

Следуя одной из основных теорий сознания — теории глобального рабочего пространства, — на этом можно ставить точку[36]. У нас есть полный комплект данных об опыте сознательного переживания яблока. Я вернусь к этой теории позже, когда буду сравнивать теорию схемы внимания с другими подходами и выдвину предположение, что многие из основных теорий на самом деле не конкурируют, а глубоко связаны друг с другом. В этой главе я хочу коротко осветить теорию глобального рабочего пространства, потому что она дает удобное представление о процессах внимания в коре головного мозга.

Согласно этой теории, информация просачивается сквозь корковую систему. Внимание отбирает и усиливает ее, пока она не превысит определенный порог и не начнет воздействовать на обширные распределенные нейронные сети в коре. В данный момент информация победила на высшем уровне — в Супербоуле соревнований по вниманию в мозге. Войдя в то состояние, при котором она может влиять на распределенные сети, информация достигла глобального рабочего пространства. По этой теории, информация в глобальном рабочем пространстве и *есть* она же в сознании. Почему глобальному рабочему пространству приписывается свойство сознания — не объяснено. В известном смысле теория эта — описательная, но не объясняющая.

Чтобы показать неполноту теории глобального рабочего пространства, давайте спросим мозг (как он описан здесь) о том, что происходит в его мире в этот момент. В предыдущих главах, когда речь шла о бессловесных животных, я “пользовался” Речинатором-5000 — гипотетическим устройством, которое можно подключить к мозгу, и оно переведет информацию в речь. Здесь нам такое устройство не понадобится, но важно, чтобы мы действовали по той же логике. Речевые механизмы в человеческом мозге могут вербализовать информацию, содержащуюся в глобальном рабочем пространстве. На первый взгляд кажется очевидным, что речь определяется внутренней информацией, но тут легко ошибиться. Интуитивно нам представляется, что, когда люди говорят, они попросту выражают свои внутренние переживания. Но на самом деле речь — это вид выходного устройства для информации. Если в системе нет определенных данных, то и словами их выразить не получится.

Иными словами, если я спрошу: “Что это перед вами?”, вы сможете ответить: “Яблоко”.

Если же я буду докапываться до деталей, вам удастся выдать и их — цвет, текстуру, форму, расположение. Вся эта информация есть у вас в глобальном рабочем пространстве.

Но предположим, я поинтересуюсь: “Осознаете ли вы яблоко? Переживаете ли его субъективно?” Тут возникнет загвоздка. Рассказывая о корковом внимании, я так и не объяснил до сих пор, как/почему в глобальном рабочем пространстве или в любой другой части коры окажется информация о сознании. Там есть информация о яблоке, а значит, можно вербализовать зрительные детали, но на чем вы могли бы основать ответ на вопрос о сознательном опыте? Вы бы не поняли и самого понятия сознания.

Нам все еще не хватает важнейшей части объяснения, что же это такое. Как я сейчас поясню, необходимо сюда добавить схему внимания.

Выше я уже говорил, почему кора головного мозга в сущности является машиной внимания. Но во внимании нет смысла, если оно мечется куда попало или им полностью управляют внешние факторы — например, яркость или громкость стимулов. Ему нужна внутренняя система управления, а она не может работать без внутренней модели. Схема внимания и есть внутренняя модель: набор данных о самом процессе внимания.

Чтобы донести до вас важность схемы внимания, я воспользуюсь аналогией, которая вначале может показаться натянутой, но вскоре вы увидите, что она работает. Я вспоминаю яркую сцену из фильма “Бутч Кэссиди и Сандэнс Кид”. Кид (в исполнении Роберта

Редфорда) претендует на роль помощника главаря банды. В качестве испытания ему нужно попасть в цель, стоя с пистолетом в руке неподвижно и прямо — это считается признаком хорошего стрелка. Однако Кид не может поразить цель и раздраженно спрашивает, нельзя ли ему пошевелиться. И тогда, в одно мгновение присев на колено и выхватив пистолет из кобуры, он делает два выстрела, каждый из которых оказывается точным. “Мне бы лучше в движении”, — небрежно замечает он.

Эта сцена — прекрасная иллюстрация инженерии управления. Лучше оставить большинство степеней свободы “свисать незакрепленными”, если это возможно. Не стоит пытаться отслеживать и контролировать всё. Вам не нужна модель во всех подробностях. Хороший управляющий механизм наблюдает, моделирует и ограничивает самое главное — в случае Сандэнса Кида это попадание пули в цель.

Другой классический пример инженерии управления — забивание гвоздя. После небольшой тренировки людям удается, размахнувшись молотком, стабильно попадать по его шляпке. Но как именно мы управляем молотком? Один из возможных подходов — отслеживать и контролировать все степени свободы. Плечо вращается, локоть распрямляется, запястье ударяет. Если оптимизировать каждое из этих движений и довести его до стойкого совершенства, возможно, вы станете блестящим забивальщиком. Но обычно, оказывается, поступают не так. Вместо этого мы сосредоточиваемся на головке молотка в момент удара по гвоздю[37]. Мы следим лишь за этой степенью свободы, потому что только этот результат имеет значение. Если заснять на видео, как люди забивают гвозди, можно заметить, что плечи

вращаются по-разному, локти колеблются, движение запястья каждый раз иное, — и тем не менее молоток умудряется снова и снова ударять по шляпке (по крайней мере, после того как вы натренируетесь как следует). Системе управления не нужно моделировать или напрямую ограничивать подробности движения, поскольку она может управлять главным результатом — в данном случае попаданием по гвоздю. Детали могут варьировать все то время, пока совокупно они приводят к нужному результату.

Стрельба Сандэнса Кида и тонкое искусство забивания гвоздей — очевидные примеры использования мозгом внутренних моделей. Мозгу нужно смоделировать суть, а не мелкие подробности. Иначе система управления даст сбой и не сможет выполнить задание — как Сандэнс Кид, которого заставили отслеживать и ограничивать слишком много несущественных деталей. Когда мозг управляет вниманием, направляя его сюда, удерживая его там, смещая медленно или быстро, расширяя или заостряя его, — какой внутренней моделью он пользуется? Какие аспекты внимания воссоздает эта модель?

В случае явного внимания, которое я описал в предыдущей главе, мозг может отслеживать его, наблюдая за определенными частями тела, — скажем, глазными яблоками, которые вращаются и направляются на различные предметы. Но в скрытом внимании не участвуют части тела. Когда скрытое внимание смещается с одного предмета на другой — с яблока на звук, а с него на воспоминание, — не происходит физического движения в пространстве. Но неуловимым образом меняется состояние миллиардов нейронов. На микроскопическом уровне скрытое внимание является тоже физическим процессом, но практичная схема внимания не станет моделировать

подробности отдельных нейронов, тормозящих связей, раундов соревнований на выбывание, корковой иерархии или нейронных сетей, распределенных по теменной и лобной долям. Мозгу не нужно знать все мелкие внутренние механистические детали, описанные ранее в этой главе. Чтобы приносить пользу, схема внимания должна отображать нечто упрощенное, сводящееся к практической сущности скрытого внимания.

Вот, наконец, мы и дошли до центрального положения теории. Мы с коллегами предполагаем, что корковая схема внимания имеет определенный вид. Информация, содержащаяся в ней, дает карикатурный пересказ того, как высшие формы коркового внимания завладевают своими предметами. Нет простого физического движения глазного яблока, как в случае явного внимания. Этот карикатурный пересказ описывает сущность, у которой нет конкретного физического воплощения, но есть расположение — где-то внутри вас, там, где она может временно завладевать объектами (яблоками, звуками, мыслями и воспоминаниями), — и которая беспрестанно движется в поиске, захватывая одни объекты и бросая другие. Когда упомянутая эфирная ментальная сущность берется за что-то, она делает это что-то ясным, видимым, несомненно присутствующим для вас, — другими словами, превращает его в опыт, переживание. И от нее есть мощный эффект. Эта сущность позволяет вам понять объект, отозваться на него, поговорить о нем или запомнить его, чтобы иметь возможность действовать в связи с ним позже. Она *дает полномочия* реагировать.

Эта аморфная сила внутри вас — лишенный подробностей отчет о корковом внимании. Не низкоуровневом, вроде состязания зрительных

сигналов в V1, а высшего его уровня, на котором предмет, например яблоко, может победить в корковом Супербоуле и повлиять на поведение.

Когда я спрашиваю: “Каково ваше мысленное отношение к этому яблоку?”, ваши механизмы вербализации могут обратиться к информации, доступной в ваших корковых сетях, — тем сведениям, что достигли глобального рабочего пространства. Я имею в виду не только зрительную информацию о яблоке, но и ту, что находится в схеме внимания и сообщает об аморфной силе внутри вас. Эти два комплекса информации связаны в единое целое, типа полного досье на яблоко в данный конкретный момент. Основываясь на обобщенных данных, вы можете сказать: “Глядя на это яблоко, я мысленно присваиваю, *сознательно переживаю* его красноту”.

Предположим, я спрошу: “Но что такое сознательное переживание? Каковы его характерные физические свойства?” Вам будет непросто ответить на подобные вопросы. Схема внимания, лишенная подробностей, не содержит никакого описания физических свойств внимания как такового.

Я продолжу вопрошать: “Можете ли вы поцарапать ваше сознательное переживание, чтобы измерить его твердость? Определить его вес? Нагреть и измерить его температуру сгорания? Какие физические измерения вообще можно с ним провести?” Вы, наверное, ответите: “У него нет никаких из перечисленных физических свойств. В каком-то смысле оно нефизическое, или метафизическое. Это просто мысленный опыт — способ моего разума что-то постичь. А вы не знаете, что такое сознание?”

Логически мозг не может сделать утверждение, если в нем не содержится информации, на которой это

утверждение основано. Теория схемы внимания обращается к комплексу информации, на котором основано утверждение о субъективном опыте. Поскольку ваш мозг строит схематическую модель коркового внимания, вы знаете, что такое сознание, и считаете, что оно у вас есть. Вы можете отвечать на вопросы о нем, а читая про него книгу вроде этой, вы более-менее представляете себе, о каком свойстве идет речь. Без схемы внимания у вас не было бы информации, необходимой для любого из этих действий, и вы бы не имели понятия, что такое “сознание”.

По моему мнению, теория сознания как схемы внимания несет в себе некоторую неизбежность. Во-первых, мы знаем, что кора пользуется скрытым вниманием. Во-вторых, мы знаем, что ей нужно управлять этим вниманием. В-третьих, мы знаем, что мозгу необходима внутренняя модель внимания, чтобы им управлять. В-четвертых, мы знаем, что подробная, точная внутренняя модель в лучшем случае бесполезна, в худшем же — вредна для работы, поэтому внутренней модели внимания придется отказаться от механистических подробностей. Следовательно, и это в-пятых, схема внимания будет отображать личность как содержащую некую аморфную, нефизическую внутреннюю силу — способность знать, испытывать и реагировать: блуждающий мысленный фокус, суть скрытого внимания без мелких подробностей. Исходя из изначальных утверждений, получается, что, если бы вам нужно было построить исправно работающий мозг с мощным корковым скрытым вниманием, вы бы сконструировали машину, которая, пользуясь созданной в ней информацией, утверждала бы, что обладает нефизическим сознанием.

Эта корковая машина, конечно, не знала бы, что ее субъективный сознательный опыт — конструкт или упрощение. Она бы приняла нефизическую природу сознательного опыта за реальность, поскольку — простите за тавтологию — мозг знает лишь то, что он знает. Он заложник своей информации.

Поразмышляйте обо всех описанных в предыдущих трех главах аспектах функционирования мозга: самих по себе их недостаточно, чтобы объяснить сознание. Сложность — это еще не все. Самый сложный в мире мозг может не содержать информации, относящейся к сознанию. И явного внимания, управляемого тектумом лягушки, недостаточно. Лягушка может поворачивать голову в сторону мухи или хищника, но не иметь в мозге никакой информации о сознании. Но недостаточно и даже более изощренного коркового внимания. Когда информация о яблоке побеждает в корковом соревновании и достигает глобального рабочего пространства, само по себе это состояние не дает сцепления с сознанием. По-прежнему нет логических причин, почему такая машина должна что-то знать о сознании, или считать, что у нее оно есть, или делать о нем утверждения. Нужно еще кое-что. А именно — внутренняя модель, описывающая корковое внимание. Схема внимания выглядит как просто крошечное дополнение, но только с ней у системы есть необходимая информация, чтобы утверждать наличие субъективного опыта. Когда в картине появляется схема внимания, сознание наконец обретает смысл для корковой машины.

Все хотят знать, у каких животных есть сознание. Выше я рассказывал, что отдельные компоненты сознания, возможно, существовали уже полмиллиарда лет назад. На рис. 4.2 показано, как эти компоненты могли в ходе эволюции собираться воедино. Если

теория схемы внимания верна, то сознание (в том виде, в котором мы, люди, его понимаем), вероятно, появилось уже давно — примерно 300 млн лет назад. Необходимые для него мозговые структуры начали возникать у рептилий, скорее всего, они есть у птиц и уж точно — у млекопитающих. Эти три группы могут различаться по сложности и разнообразию скрытого внимания и схем внимания, но у всех у них в том или ином виде может наличествовать то, что мы зовем сознанием.



Рис. 4.2. Вероятная эволюция сознания от губок до млекопитающих

Но я не закончил с эволюцией. Самая человеческая часть еще впереди. Мы не только строим богатые описательные модели самих себя, но и безотчетно приписываем сознание друг другу, создавая социальную экосистему. Мы видим сознание в других людях, домашних животных, игрушках, в обширном и невидимом мире богов и духов, которых проецируем в окружающее нас пространство. В следующей главе я обращусь к этому важнейшему социальному применению сознания, которое так масштабно расширилось у нашего вида.

Глава

5

Социальное сознание

Нам, людям, действительно дано догадываться о душевном состоянии друг друга. Но при этом мы не занимаемся какими-то намеренными наблюдениями, которые можно свести воедино для интеллектуального понимания, что происходит с другими. (Впрочем, иногда мы пытаемся это делать, но не достигаем особых успехов.) Взамен этого у нас есть тщательно настроенная интуиция. Мы словно *знаем*, что думают и чувствуют другие люди. Иногда это знание так осязаемо, что мы будто напрямую ощущаем мысли и чувства других как некое излучение. Конечно, на самом деле этого не происходит. Но миллионы лет эволюции подарили нам возможность считывать тонкие намеки и строить подробные модели психического состояния друг друга, — причем делаем мы это скорее интуитивно, нежели явно.

Мы приписываем друг другу весь спектр содержаний психического мира: эмоции, намерения, интересы, убеждения. Блистательный в своей сложности процесс воссоздания чужого внутреннего мира называется построением *модели психического*[9], или *модели психического состояния*, Theory of Mind[1]. Это не теория в интеллектуальном смысле, это процесс, происходящий автоматически,

неизбежно, — мы просто не можем этого не делать. Но в построении моделей психики друг друга особенно важен один компонент: воссоздание чужого внимания.

Как узнать, потянетесь ли вы за этим яблоком, если неизвестно, обратили ли вы на него внимание? И даже если знать, что вы уделили яблоку внимание, разве можно предсказать, что вы дальше скажете или сделаете, если не понимать последствий этого внимания? Мой первый шаг в воссоздании вашего психического мира — понять, что психические процессы могут быть на чем-то сфокусированы, что этот фокус в зависимости от обстоятельств может быть узким или широким, он в состоянии перемещаться от предмета к предмету, у него есть предсказуемые последствия. Без этого никакой модели психического состояния не получится. Мне нужно не просто построить модель *содержания* психики, мне необходима модель того, что *представляет собой* эта психика.

Если я стою напротив вас, то вижу, куда вы смотрите. Целая научная область выросла вокруг того, как мозг обрабатывает направление чужого взгляда[2]. Но для воссоздания вашего внутреннего мира мне недостаточно знать, куда направлено ваше явное внимание. Еще нужно разобраться со *скрытым*. Я должен уловить все возможные подсказки из контекста, в который входят ваши движения, выражение лица, слова и мое знание вас в целом. Не так важно, куда направлен ваш взгляд и видны ли мне ваши глаза, — нужно воссоздать информацию, которая просочилась по вашей корковой иерархии и достигла высших уровней обработки, и понять, как

эти высшие уровни обработки могут повлиять на ваше поведение.

Ручаюсь, что никому и никогда при взгляде на другого человека не приходила в голову немедленно интуитивная мысль типа такой: “Корковые зрительные пути моего визави в данный момент вовлечены в обработку множественных стимулов; нейроны, отражающие форму яблока, развили повышенную активность, реагируя на сигналы из зон лобной доли; как следствие этого, повышенная активность нейронов частично затормозила соседние нейронные репрезентации путем латерального торможения, основанного на локальных промежуточных нейронах, которые используют гамма-аминомасляную кислоту в качестве нейромедиатора…” Можно и дальше продолжать в том же духе. Но на самом деле никто никогда не приписывает другому человеку истинное, физиологическое, нейронное внимание. Нам это не нужно, особенно в таких подробностях. Вместо этого мой мозг строит намного более схематическую и эффективную модель. Я интуитивно понимаю: “Прямо сейчас сознание этого человека воспринимает яблоко, у чего может быть много разных последствий”.

Мы снова возвращаемся к сознанию как к упрощенной, практичной модели внимания. Но теперь схема внимания используется в режиме социального интеллекта: мы моделируем не себя, а другого человека.

Салли и Энн пришли в парк с двумя закрытыми корзинами для пикника и расположились на отдых. Через некоторое время Салли убрала свой бутерброд в корзину А и отошла в туалет. Пока ее не было, Энн тайком переложила бутерброд в корзину Б и закрыла

крышки обратно. В какую корзину сначала заглянет вернувшаяся Салли в поисках своего бутерброда? Этим простым вопросом привыкли проверять способность человека к построению модели психического состояния другого[3].

Чтобы пройти этот тест, нужно учитывать сведения, имеющиеся во внутреннем мире Салли. Ее знание о расположении бутерброда верно, когда она кладет его в корзину А, но становится ложным, когда бутерброд перекладывают в корзину Б. Эту задачу невозможно решить без представления о психическом мире Салли как об отдельной сущности, которая содержит информацию (возможно, ложную), определяющую действия девушки. Из-за этого тест иногда называют задачей на понимание ложных убеждений. Правильный ответ: она заглянет в корзинку А и увидит, что ее бутерброд пропал.

Дети младше пяти лет далеко не всегда дают правильный ответ[4]. В их представлении, если бутерброд в корзине Б, то там Салли и должна его искать. Зачем ей открывать корзину, в которой бутерброда нет? Когда дети перешагивают пятилетний рубеж, их социальное мышление настраивается и задача оказывается для них интуитивно понятной. Когда мы становимся взрослыми, нам обычно уже неплохо удается следить за психическими состояниями других людей.

Некоторые успехи в решении задачи на понимание ложных убеждений показывают и шимпанзе[5]. Разыграем перед ними сценарий Энн и Салли. Салли кладет фрукт в коробку, уходит, Энн перекладывает его в другую коробку. Салли возвращается забрать фрукт. По движениям глаз шимпанзе видно, что они больше смотрят на коробку А (в которой оставила

фрукт Салли и куда она скорее всего заглянет, вернувшись), чем на коробку Б, в которой на самом деле находится фрукт. Судя по всему, шимпанзе учитывают содержание психики Салли и предвосхищают ее действия.

Похоже, что вороны тоже способны к решению подобных задач[6]. Они часто прячут пищу, но не любят, когда другие птицы крадут сохраненные ими лакомства. Вот одна птица прячет вкусненькое на глазах у другой. Затем подсматривающая птица улетает. И тогда первая ворона аккуратно перепрятывает еду — наверное, чтобы ее потом не украли. Запасливая птица будто понимает, что вторая ворона подсмотрела, где спрятана еда, — так что, когда она вернется, ее убеждение окажется ложным и она будет искать не там.

В принципе решить задачу на понимание ложных убеждений способны немногие из нечеловекоподобных животных. Даже описанные выше исключения вызывают вопросы. Но, по-моему, было бы преждевременно делать вывод, что у остальных животных отсутствует модель психического состояния. Дело в том, что задача на понимание ложных убеждений — слишком высокая планка. Наблюдать за несколькими коробками с разнообразным содержимым — интеллектуально сложное задание, сравнимое с игрой в наперстки. Неудивительно, что лишь люди способны раз за разом с этим справляться. Меня интересует кое-что попроще: понятие о внутреннем мире. Мы знаем, что у Салли есть внутренний психический мир, а внутренний мир — это нечто, в чем содержится информация и что будет управлять поведением на основании этой информации. Есть ли у других

животных такое же интуитивное понимание? Знают ли они, что для другого значит “осознать что-то”?

Ученые, которые изучают поведение животных, предпочитают простые объяснения. Вместо допущения, что у животного есть представление о чужом сознании, проще предположить, что животное попросту выучило набор несложных правил. К примеру, зебре необязательно знать, что ее заметил лев. Ей достаточно просто убегать от всего большого и зубастого. Стимул на входе, реакция на выходе. Если у зебры накоплена достаточно большая база таких ассоциаций, ей удастся выжить. Стоит, однако, отметить, что подобная гипотеза, столь типичная для психологии “стимул-реакция”, на самом деле весьма наивна. Обширная база выученных ассоциаций — не самый простой и не самый эффективный метод ориентирования в сложной среде. С вычислительной точки зрения подход, основанный на моделях, был бы проще, поскольку одна модель может обслуживать большое разнообразие обстоятельств. Для зебры, возможно, окажется проще и вычислительно дешевле построить схематическую модель, в которой у льва есть психическое содержание, туда попадают предметы из окружающего мира, а когда это случается, модели психического мира могут управлять поведением хищника.

Предположение, что зебра “понимает” сознание другого животного, кажется нам неправдоподобным, но это лишь потому, что мы считаем сознание благородной характеристикой — связанной с культурой, присущей исключительно людям. Зебрам не хватает сложности и поэтичности. Но подобные мысли подсовывает нам эго. Я считаю, что сознание — древняя составляющая модели психического состояния, простая и эффективная модель, нужная,

чтобы предугадывать поведение других животных, и, скорее всего, она развилась задолго до человечества. Не удивлюсь, если зебры, прочие млекопитающие, птицы и, может быть, даже некоторые рептилии используют этот удобный конструкт — сознание (разной сложности), чтобы предугадать, как себя поведет другой.

Излюбленный стереотип — считать, что высшим сознанием обладаем именно мы, люди. Мы полагаем, что у других животных сознания вообще нет или оно менее развито. Такая точка зрения согласуется с распространенным предположением, что сознание возникает от сложности. Раз среди всех животных именно у человека самые сложные мозги, значит, у него должно быть и самое лучшее сознание. Но из тех умственных талантов, которыми мы любим похваляться, — язык, математика, использование инструментов и т.д. — сознание, пожалуй, самый примитивный и наименее выделяющий нас среди других. Готов признать, что содержание сознания — мысли, идеи, убеждения, озарения, знание о смерти — у людей, вероятно, сложнее, чем у других животных. Но сам факт наличия сознания, способности переживать субъективный опыт чего-либо и приписывать подобный опыт другим — настолько базовая необходимость, что ее могут разделять с нами многие другие представители животного царства. Если теория схемы внимания верна, то сознание точно есть не только у людей.

Если я хочу в чем-то разобраться получше, то обычно нахожу полезным использование инженерного подхода. Давайте продумаем, как бы мы конструировали машину, предсказывающую человеческое поведение в каждый конкретный момент. Это упражнение покажет нам, насколько

трудную задачу решает мозг, и в то же время раскроет принципы, лежащие в основе его работы. Оно также укажет направление к дальнейшим главам этой книги, в которых я буду рассуждать, в состоянии ли мы построить искусственное сознание.

Допустим, человек (назовем его Кевином) входит в комнату. Скрытая камера наблюдает за ним, а микрофон подслушивает. Камера и микрофон подключены к компьютерной системе, Предсказателю-5000, задача которой — все время выдавать предсказания того, что Кевин сейчас сделает.

Вот что есть в комнате: на столе в ее центре — пончик, посыпанный сахарной пудрой, ярко освещенный лампой с потолка; лужица воды на полу перед столом; мобильный телефон на полке в углу, где мало света.

Первое задание, которое мы дадим предсказательному устройству, — определить в комнате аффордансы. Термин “аффóрданс” ввел психолог и натуралист Джеймс Гибсон в 1970-х гг.[7] Он понял, что, когда человек или животное смотрят на мир, настоящая задача их зрительной системы — не вобрать в себя мир таким, какой он есть (будто фотографируя), а определить возможности для действий. Подобные возможности он назвал аффордансами. Муха создает аффорданс для лягушки: муху можно поймать и съесть. Ветка дает аффорданс птице: на ветку можно сесть. Дверная ручка — аффорданс для человека: ее можно повернуть.

Сразу видна колоссальная сложность задачи — построить работающий механизм предсказания поведения. Предсказателю-5000 нужно знание человеческого поведения, чтобы увидеть человека,

увидеть лужу — и извлечь информацию о том, что тот способен переступить через лужу. Представьте, что эта задача повторяется для каждого объекта, который может встретиться человеку. Эффективное устройство предсказания должно иметь обширные знания о повадках людей. Хотя такое знание в принципе не слишком сложно воплотить в искусственном интеллекте (это вопрос распознавания объектов и ассоциирования их с возможными действиями), но нужно же выучить огромное количество ассоциаций.

Чтобы дополнительно усложнить Предсказателю-5000 задачу, добавлю, что у каждого объекта может быть много аффордансов. Обстановка комнаты на первый взгляд кажется спартанской — всего-то три предмета, — но в ней скрывается огромный набор возможностей. Кевин может переступить через лужу, шлепнуть по ней (чтобы разбрызгать воду), достать платок и вытереть ее. Он может в гневе перевернуть стол, или аккуратно передвинуть его в другую часть комнаты, или залезть на него, или спрятаться под ним. Он может взять пончик и съесть его, или же бросить на пол и растоптать, либо, забавляясь, приложить к глазу и посмотреть сквозь дырку, как в монокль. Он может взять телефон и попробовать позвонить, или исподтишка припрятать телефон в карман, или просто подойти к телефону и рассматривать его. Механизм предсказания сталкивается с огромным количеством возможных аффордансов.

Упростим задачу и дадим Предсказателю-5000 некоторую информацию об априорной вероятности — общее знание о том, как ведет себя среднестатистический человек. Большинство людей не шлепают по лужам, не переворачивают столы и не топчут пончики. Это события низкой вероятности.

Намного выше вероятность того, что человек переступит через лужу и съест пончик. Имея достаточно данных о поведении людей, мы можем вычислить вероятность того, что “усредненный” человек поступит так-то или так-то, встретив на пути лужу или увидев пончик.

Но предсказательный механизм не заработает даже после того, как мы загрузим в машину все эти справочные сведения. Информация, имеющаяся у нас, описывает стандартного, обычного человека. Мы не знаем, как поведет себя в данном случае конкретный экземпляр, Кевин. Давайте еще поможем Предсказателю-5000 и дадим ему совокупный инструментарий традиционной модели психического состояния. Например, если мы знаем, что Кевин десять часов не ел, можно предположить повышенную вероятность пожирания пончика. Если знать, что у него диабет, можно предположить, что пончик он есть не станет. Если известно, что у Кевина проблемы с самоконтролем или он сильно рассержен, можно предположить, что повышена вероятность топанья по луже. Уже упоминавшийся философ Дэниел Деннет называет такие знания о человеке интенциональными установками[8]. Глядя на других, мы всегда автоматически задаем себе вопрос: “Могу ли я разгадать мотивацию этого человека, чтобы предсказать, что он будет делать дальше?”

В последнее время исследователи стали строить искусственные системы в попытке решить задачу предсказания поведения, используя интенциональные установки[9]. Чтобы угадать чью-то такую установку, предположительно требуется накопить затейливые знания о человеческой природе — о типичных мотивациях и особенно о выражении эмоций через

мимику. Однако я хочу отойти от интенциональных установок и обратить внимание на кое-что попроще, но, как мне кажется, принципиально более важное.

Предположим, я возьму доступную статистическую информацию о человеческом поведении в целом, добавлю к ней сведения о конкретном Кевине и его особой интенциональной установке на сегодняшнее утро и все это загружу в общий набор цифр. Тогда можно будет вычислить вероятность реализации каждого аффорданса в комнате. Съесть пончик: 30%. Осторожно переступить через лужу: 50%. Забрать чей-то забытый на полке телефон: 3%. Я проделал всю подготовительную работу, которая обычно считается сущностью социального познания, и выдал эту информацию Предсказателю-5000. Кажется, я выполнил за него его должностные обязанности.

Тем не менее даже со всей этой загруженной в нее полезной информацией машина все еще не в силах предсказать поведение Кевина в каждый конкретный момент. Ей нужны данные по определяющей скрытой переменной, которая варьирует во времени: куда Кевин направляет внимание. Ресурсы Кевина по обработке данных постоянно перемещаются, двигаются по комнате. В результате этого вероятности действий с пончиком, лужей и телефоном постоянно меняются.

Вот, скажем, пончик. Априорная вероятность того, что его бросят на пол и растопчут, настолько мала, что предсказательный механизм отбрасывает этот вариант. Так же как и вариант с пончиком в роли монокля. Нам остается лишь один реалистичный вероятный аффорданс: пончик съедят. Но вероятность этого действия меняется во времени. Чтобы попытаться ее предсказать, понадобится немножко

математики: я воспользуюсь системой, которая называется байесовской статистикой. Если вы математикой не интересуетесь, пропустите эту часть, но я обещаю объяснить как можно проще.

Предположим, что, с учетом всех наших знаний о людях в целом и о Кевине в частности, мы считаем вероятность поедания пончика равной примерно 30%. Это число называется априорной вероятностью: именно его мы загрузили в нашего Предсказателя. Обозначим это число $P_{\text{априори}}$. Теперь предположим, что машина высчитала второе число, которое тоже лежит между 0 и 100%, — примерную оценку объема внимания, которое Кевин направляет на пончик. Назовем это число C_1, где индекс 1 означает объект 1 — пончик. Позже мы рассмотрим C_2 и C_3, говоря об объеме внимания, которое Кевин обратит на лужу и телефон. C_1 постоянно меняется во времени. Бо́льшую его часть Кевин обращает на пончик мало внимания и значение C_1 близко к нулю. Иногда внимание к пончику может усилиться, и C_1 чуть возрастет; или же почти все внимание обратится к пончику, и тогда C_1 временно окажется близко к отметке 100%.

Чем больше внимания Кевин обращает на пончик (т.е. чем выше значение C_1), тем больше вероятность, что он с ним как-то поступит. C_1 — что-то вроде разрешительной переменной, которая открывает возможности к действию. Сейчас будет единственное уравнение, которым я вас нагружу. Допустим, вы спросили Предсказатель-5000: “Какова вероятность в этот конкретный момент времени (назовем ее $P_{\text{действия}}$), что Кевин съест пончик?”

Машина оценивает эту вероятность с помощью простого вычисления:

$$P_{\text{действия}} = C_1 \times P_{\text{априори}}.$$

Вот и все. Умножьте априорную вероятность на предполагаемый объем внимания — и вы сможете предсказать поведение Кевина в каждый конкретный момент. Пока он не уделяет внимания пончику, $C_1 = 0$, а значит, и $P_{\text{действия}} = 0$, и машина предсказывает, что сейчас есть он не будет. По мере того как внимание к пончику возрастает, возрастает и оценочная вероятность того, что Кевин его съест. Но даже на пике внимания вероятность поедания лакомства никогда не превысит $P_{\text{априори}}$, которая может быть не так уж и велика, — просто потому, что люди в принципе не так часто хватают и съедают случайно попавшиеся им на глаза пончики. Когда внимание к пончику снова падает, стремится к нулю и вероятность его поедания. Удобство таких вычислений в том, что они применяют стандартный подход к построению модели психического состояния (который требует рассматривать один за другим статичные моменты во времени: если дать Кевину пончик, съест ли он его?) и помещают его в систему, которая способна учитывать динамические изменения внимания во времени.

Задача машины — оценивать постоянно меняющееся число C_1. Но наше искусственное предсказательное устройство не имеет прямого доступа к мозгу Кевина. Внимание исследуемого — результат ряда сложнейших нейронных взаимодействий, спрятанных в его черепной коробке. Машина всего лишь направляет на Кевина камеру и микрофон. Ей нужны какие-то простые правила, которые могут превратить данные ограниченных наблюдений в оценку состояния внимания Кевина.

Чтобы помочь нашей машине, привлечем некоторые хорошо обоснованные научные данные о том, как работает внимание. Во-первых, пончик белого цвета и расположен под яркой лампой: это стимул перцептивно значимый, он бросается в глаза. Пончик высокой насыщенности означает, что Кевин с большей вероятностью обратит на него внимание. Следовательно, машина может повысить оценочную величину C_1. Во-вторых, пончик — единственный предмет на столе, а внимание находится в обратной зависимости от зашумленности, или зрительной конкуренции. На основе этой подсказки машина опять же может повысить оценку C_1. В-третьих, полезную информацию дает направление взгляда Кевина. Взгляд — не точный индикатор, Кевин может смотреть прямо на пончик, но думать о чем-то другом — скажем, прислушиваться к шуму из коридора или строить планы на завтра. Но в общем и целом, если вы хотите оценить чье-то внимание, направление взгляда стоит учитывать. В-четвертых, подсказкой способно оказаться выражение лица. Если взгляд останавливается на пончике и лицо быстро меняет выражение с нейтрального на более сосредоточенное, — у машины есть все основания, чтобы вычислить резкое повышение значения C_1.

Учитывая все эти данные — взаимосвязь внимания с насыщенностью стимула, зашумленностью среды, направлением взгляда и выражением лица, — машина может оценить меняющееся во времени значение C_1, объем ресурсов обработки, которые Кевин направляет на объект 1. Эта оценка меняется во времени по мере того, как наше устройство получает новые сведения. Используя значение C_1, машина может для каждого момента времени вычислить вероятность того, что

Кевин съест пончик: сейчас — может быть; сейчас — нет; точно нет; нет; сейчас — опять готов.

История становится еще интереснее, когда мы вспоминаем, что в комнате есть и другие объекты. Скажем, лужа на полу. Машина вычисляет значение C_2, отражающее оценку объема внимания, которое Кевин уделяет объекту 2. Лужа находится не в самой очевидной позиции. Она в тени стола, поэтому не бликует на свету. Другими словами, этот стимул не перцептивно значимый. Предположим, Кевин не направит взгляд на пол. Пока еще остается вероятность, что он заметит лужу периферическим зрением, но в первом приближении Предсказатель может обоснованно оценить значение C_2 как низкое — т.е. малую вероятность того, что Кевин перешагнет через лужу по дороге к столу. Если он не обратит на нее внимания, то прямо туда и угодит.

Два упомянутых выше числа, C_1 и C_2, зависят друг от друга. Они взаимосвязаны. Ресурсы Кевина ограниченны, поэтому с увеличением C_1 должно уменьшаться C_2. Когда он обращает больше внимания на пончик, меньше внимания достается луже, и наоборот. Если его внимание приковано к пончику, то, скорее всего, по дороге к столу он наступит в лужу. Следовательно, вычисление C для каждого объекта в комнате требует более сложной модели, учитывающей конкурентные взаимосвязи между объектами.

Также машина может вычислить и C_3, приблизительную оценку того, сколько внимания Кевин уделит третьему объекту — телефону. Изначально машина выдает низкое значение C_3, поскольку телефон спрятан в тени в дальнем углу комнаты и Кевин на него не смотрит. Теперь телефон издает звонок, обретая перцептивную значимость.

Отметив при помощи микрофона это изменение насыщенности, наш предсказательный механизм может вычислить резкое повышение C_3. Внимание Кевина в этот момент, скорее всего, метнулось к телефону. Даже если его взгляд направлен не туда, при такой насыщенности стимула в данный момент значение C_3 должно быть велико. Как прямое следствие этого, C_1 и C_2 должны в тот же миг резко снизиться. Три числа пребывают в постоянном соревновательном танце друг с другом. В момент телефонного звонка вероятность поедания пончика быстро падает до нуля. Но вот что еще важно: внимание Кевина, которое привлек зазвонивший телефон, обладает некоторой инертностью. У него будет наблюдаться тенденция оставаться направленным на телефон по крайней мере еще полсекунды после звонка и лишь затем постепенно снижаться по кривой. И это типично для человека. Внимание не сдвигается с предмета на предмет в мгновение ока. Чтобы наилучшим образом предсказать поведение, при вычислении C нужно учитывать приблизительную инертность человеческого внимания.

Нашему Предсказателю-5000 придется вычислять постоянно меняющиеся значения C_1, C_2 и C_3 с совокупным учетом всех перечисленных сложностей. Вычисления основаны на базовых знаниях о том, как работает человеческое внимание. Машина пользуется данными о перцептивной значимости различных стимулов в окружающей среде, о ее зашумленности, о направлении взгляда Кевина, выражении его лица и временно́й динамике человеческого внимания. Вычисляя меняющиеся во времени значения C_1, C_2 и C_3, наше устройство сможет оценивать постоянно

меняющуюся вероятность совершения Кевином действий, аффордансы для которых дают пончик, лужа и телефон.

Подобное вычисление *С* равносильно модели. Предсказательное устройство строит описание, в котором Кевин обладает неким свойством — назовем его субстанцией *С*. Субстанция невидима, ее невозможно наблюдать напрямую. Она не отражает и не поглощает свет. Субстанция зарождается внутри Кевина и течет вовне. Для нее прослеживается тенденция идти по прямым линиям, исходящим из глаз, хотя иногда ее траектория не подчиняется этому правилу, двигаясь в направлениях, не связанных со зрением. Эта субстанция взаимодействует с определенными предметами в среде. Она может разделяться по разным объектам, но только в конкурентном порядке, т.е. сильнее сосредоточиваясь на одном, она уменьшает свое присутствие на других. Словно вода из лейки: чем больше Кевин поливает ею один предмет, тем меньше остается на другие.

Субстанция *С* во многом похожа на классическую жидкость[10]. У нее есть источник. Ее объем в целом ограничен — я имею в виду, что увеличение потока в одном направлении означает его уменьшение в другом. Субстанция обладает некоторой инертностью — направление она меняет с ленцой. Она напоминает невидимую вязкую жидкость, но в то же время обладает энергией или волей — в том плане, что дает их действующему субъекту. Субстанция не побуждает Кевина действовать напрямую и не определяет конкретных его действий. Ее присутствие работает как заряжающее энергией поле, которое дает Кевину ресурсы для поведенческого выбора.

Субстанция *С* — конструкт, созданный нашим предсказательным устройством. На самом деле как таковой ее не существует. Это заменитель для обозначения настоящей, предельно сложной нейронной обработки информации, происходящей в голове Кевина. Его действиями управляют примерно 8 млрд нейронов в его мозге, а Предсказатель приписывает ему нечто намного более простое, карикатурное, напоминающее метафизическое сознание.

Субстанция *С* — упрощенная, модельная версия коркового внимания. Это схема внимания.

Для подобного предсказательного устройства в перспективе открываются малоприятные прикладные применения. Например, его можно устанавливать в магазинах, чтобы предсказывать поведение покупателей. Или использовать для безопасности и управления толпами, отслеживая внимание людей на десять шагов глубже, чем позволило бы простое наблюдение, кто где находится и куда смотрит. Возможно его применение в компьютерных играх, чтобы запрограммированные в игре персонажи лучше предсказывали поведение реальных игроков. Приписывание действующим субъектам невидимой метафизической силы сознания — не пафосное поэтическое преувеличение, а простой программируемый прием, полезный для предсказания поведения.

Наш Предсказатель-5000 вряд ли достигнет особых успехов. Человеческое поведение сложно предвосхитить хоть с какой-то точностью. Кевин может просто бродить по комнате, ничего не предпринимая и бормоча себе под нос, — такой поток поведения не очень-то поддается предсказанию.

Реакции Кевина могут быть хаотическими или определяться неведомыми нам скрытыми факторами. Не думаю, что люди так уж преуспевают в предугадывании поведения друг друга в каждый момент времени, и не думаю, что в нем преуспеет искусственный интеллект. Но машине и необязательно быть абсолютно успешной. Если ее предсказания *относительно* хороши, если они точнее случайных догадок, то они все равно будут давать полезное преимущество — подобное тому, что мы, люди, получаем от своих несовершенных социальных предсказаний.

Озадачим машину разбором с еще более трудной ситуацией. Кевин может направить внимание на мысль или эмоцию точно так же, как на пончик или телефон. Если добавить в задачу абстрактные направления внимания, она становится заметно больше. Представим, что наш Предсказатель-5000 через микрофон слушает, как Кевин говорит по телефону. Задача машины — воссоздать колеблющееся внимание Кевина, направленное на темы разговора, и предсказать, что он будет говорить дальше. Задача с телефоном очевидно сложнее, чем с видео, поскольку информация, которую нужно собрать машине, и намного сильнее урезана, и намного более абстрактна. Всё должно быть передано словами, которые произносит Кевин, и его интонациями. Но в основе лежат те же принципы. Машина определяет объекты (в данном случае абстрактные мысли, которые могут привлечь внимание Кевина), моделирует субстанцию *С* внутри Кевина (сущность осознания, которая делится между этими объектами) и затем использует эту модель, чтобы предсказать вербальное поведение Кевина.

Я полагаю, что в людях содержатся именно такие предсказательные механизмы. Мы постоянно приписываем друг другу тонкую субстанцию *С*, осознание предметов, некую силу или сущность, которая невидима и может течь, словно жидкость. Она исходит из человека, ей свойственно истекать из его глаз по прямой. Эта субстанция, или сущность, входит в контакт с предметами окружающей среды. У нее есть энергия или подобное воле свойство — в том плане, что она дает людям ресурсы делать поведенческий выбор и таким образом воздействовать на мир. Мы не всегда способны осознать, что сами строим эту причудливую модель внимания. Она возникает автоматически, создавая у нас удивительное впечатление, что мы можем воспринимать сознание других людей, исходящее из них. И все мы это делаем потому, что эволюция нашла практичное решение важной проблемы — предсказания поведения других.

Моя лаборатория недавно провела забавный эксперимент[11]. Его участники видели на мониторе изображение бумажного цилиндра, вертикально стоящего на столе. Их просили представить себе, что цилиндр постепенно наклоняется, и прикинуть, при каком угле наклона он упадет. Угол нужно было оценить, используя стрелки на клавиатуре. Проба за пробой, мы предъявляли цилиндры разной высоты и ширины. Иногда участникам предлагали представить наклон влево, иногда вправо. Люди неплохо справлялись с этим заданием, их оценки были вполне правдоподобны с физической точки зрения.

Одновременно с цилиндрами в каждой пробе на экране присутствовало изображение лица в профиль. Мы не давали участникам никаких комментариев по его поводу, оно просто висело на удалении либо в

левой, либо в правой части экрана и “глядело” на бумажный цилиндр. Когда после эксперимента мы спрашивали участников о том, зачем оно там было, они предлагали самые разные объяснения, но никто не догадался о настоящей цели эксперимента. Большинство считало, что лицо никак не повлияло на выполнение экспериментальной задачи. Тем не менее оно внесло свой небольшой вклад. Участники словно бы воспринимали взгляд как луч энергии, исходящий из глаза, направленный на цилиндр и влияющий на угол наклона, при котором тот упадет. Когда цилиндр наклонялся по направлению к лицу, луч из глаза будто подпирал его, и, по оценкам участников, цилиндр можно было наклонить чуть сильнее, прежде чем он упадет. Когда цилиндр наклонялся от лица, луч из глаза будто подталкивал его, и испытуемые оценивали, что он упадет быстрее при меньшем угле наклона. Но эффект пропадал, если у лица на экране глаза были завязаны. Предполагаемый критический угол наклона оставался неизменным независимо от того, куда клонился цилиндр — к лицу или от лица с завязанными глазами, — повязка будто бы останавливала луч из глаза и его воздействие.

По воздействию луча на кренящийся бумажный цилиндр мы даже смогли вычислить значение его воображаемой силы. Величина совсем небольшая, способная лишь слегка подтолкнуть цилиндр: около одной сотой ньютона. Для сравнения — это чуть меньше, чем сила, с которой при нормальном земном притяжении воздействует на вашу ладонь лежащая на ней изюминка. Такая ее “скромность” была, конечно, ожидаема. Если бы люди воспринимали исходящую из чужих глаз силу достаточно мощной (скажем, сравнимой с ветром, который может опрокинуть урну с мусором), конфликт между восприятием и

реальностью стал бы слишком очевиден и ошибки восприятия сильно мешали бы нам выживать. Но нет — воображаемая сила остается настолько малой, что едва поддается измерению.

Участники нашего эксперимента не осознавали это искажение восприятия, которое зато проявилось в результатах. В анкете, заполняемой после исследования, они настаивали, что лицо никак не повлияло на их оценки угла наклона. На всякий случай мы спрашивали их также о том, как, по их мнению, работает зрение: это что-то исходящее из глаз или что-то входящее в них? Около 5% испытуемых не очень разбирались в предмете и написали, что исходящее. Остальные верно указали, что механизм зрения связан с попаданием света в глаза. И, несмотря на это, безотчетно все они будто воспринимали открытый глаз как источник невидимой субстанции, которая взаимодействует с физическим миром. И не могли иначе. По-моему, мы добрались до субстанции *С*. Мы наблюдали работу упрощенной модели внимания.

Представление о том, что зрительное внимание может исходить из глаз и физически влиять на предметы окружающего мира, далеко не ново. Это одно из неизбывных убеждений народной мудрости. Оно называется теорией зрительных лучей и берет начало как минимум в древнегреческой философии[12]. Эта теория просуществовала довольно долго, и ее поддерживали такие знаменитости, как Платон, а позже — римский врач Гален. На рубеже X–XI вв. арабский ученый Ибн аль-Хайсам наконец вывел верные законы оптики и объявил теорию зрительных лучей неверной. Изображение образуется, когда свет входит в глаз по прямым линиям.

Несмотря на окончательное научное решение этого вопроса, народная вера в лучи из глаз не сдается. И через тысячу лет после аль-Хайсама во многих культурах всерьез верят в сглаз и зарабатывают кучу денег на защищающих от него амулетах[13]. В нашей культуре эту веру олицетворяет Супермен, который обладает рентгеновским зрением, каким-то образом выстреливающим из его глаз и сжигающим предметы. Почти каждый из нас испытывал неприятное ощущение, что тебе в спину смотрят, — как будто луч энергии касается затылка и шеи. В 1898 г. психолог Эдвард Титченер решил, что эти убеждения заслужили своей распространенностью экспериментальную проверку[14]. В контролируемых тестах он обнаружил, что люди не могут напрямую чувствовать взгляды друг друга — сколь бы сильно им это ни казалось.

Вера в лучи, исходящие из глаз, психологически соблазнительна. Когда мне было лет пять, я сидел на ступеньках крыльца с отцом и смотрел на звезды. Отец спросил, как так получается, что мы видим далекие-предалекие предметы. Вспоминая об этом сегодня, я думаю, что он хотел рассказать про свет из древности, который путешествует в течение миллионов лет, но, боюсь, я сбил его с мысли. Я начал объяснять, что, когда ты смотришь вверх, из твоих глаз выходит такая штука, зрение, и отправляется в небо. Я так и вижу, как ученый в нем сморщился от досады. Не теряя времени, он прямо на месте доходчиво объяснил мне основы оптики. Да, я точно помню момент, когда переключился с теории зрительных лучей — иными словами, эмиссионной теории зрения — на интромиссионную. Возможно, это был поворотный момент для моей будущей жизни ученого, хотя оказывается, что такой поворот проходит каждый человек. Вера в связь зрения с чем-то исходящим из

глаз столь интуитивно удобна, столь глубоко заложена в наше понимание активного зрительного восприятия, что дети считают так по умолчанию, согласно открытию психолога Жана Пиаже[15]. Лишь много позже нас обучают истинной научно достоверной версии.

Ряд исследований 1990-х гг. наводит на мысль, что большинство студентов американских университетов верили в ошибочную теорию зрительных лучей[16]. В этих исследованиях аж 60% студентов утверждали, что зрение происходит из-за чего-то исходящего из глаз, а не поступающего в них. Должен сказать, что я скептически отношусь к подобным радикальным заявлениям. Возможно, формулировки вопросов повлияли на ответы, или естественно-научное образование резко улучшилось за последние 30 лет, или участники просто-напросто троллили экспериментаторов. Наши данные противоречат столь значительной доле ложных убеждений, а мы исследовали довольно широкий демографический срез с разными уровнями образования и дохода обследуемых. В наше время серьезное убеждение в теории о том, что зрение — это лучи из глаз, встречается редко, большинство взрослых людей знакомо с азами оптики.

И тем не менее, несмотря на то что с научной точки зрения нам все понятно, мы по-прежнему цепляемся за мысленное представление о чем-то исходящем. Для меня самым интересным результатом исследования с бумажным цилиндром стало резкое отличие между умственным знанием и безотчетной интуицией человека. Неважно, обладали ли участники научными знаниями: на неосознаваемом уровне они все еще воспринимали зрение как нечто, идущее из

человеческих глаз. Подозреваю, что это интуитивное понимание не ограничивается зрением, а простирается на любой вид внимания. Наша культура полна наивных представлений о том, как сосредоточенное внимание может протянуться и коснуться человека или предмета. Интересно, много ли детей после первого просмотра “Звездных войн” попытались сконцентрироваться и сдвинуть карандаш силой мысли? Лично я пытался. По-моему, классическое представление о лучах из глаз имеет отношение скорее к *мысленным* лучам, которые зачастую исходят именно оттуда, но необязательно ограничиваются зрением.

Эти представления так крепко держатся в культуре, вероятно, потому, что берут начало в глубинной, автоматической, неосознаваемой модели, с которой рождается каждый и которая развивалась на протяжении миллионов лет. Такая модель позволяет нам схематично, но эффективно отслеживать внимание других людей, чтобы лучше прогнозировать их поведение. Даже если наши знания говорят нам об обратном, мы не в силах заглушить интуитивное понимание. Мы не можем не учитывать подспудно лучи внимания других людей.

Мне кажется, в этих интуитивных представлениях кроется также причина того, почему так трудно изучать сознание научно. Множество теорий сознания попадаются в ловушку ненаучного, в сущности, мистического допущения. Если мы будем пытаться уразуметь, как человеческий мозг генерирует неосязаемую субстанцию *C*, нам никогда не прийти к научному пониманию. Люди смотрят внутрь себя и исследуют модели, эволюционное предназначение которых вовсе не в том, чтобы создавать научно достоверные описания реальности. Эти модели

формировались и совершенствовались в нас последние 300 млн лет эволюции или даже больше для того, чтобы быть полезными и удобными в определенных практических целях. Они схематизированы и очищены от лишних деталей. Нас же тянет воспринимать содержание этих моделей буквально и пускаться в научные поиски исходящих из глаз лучей или незримой внутренней энергии, субъективной сущности. Мы пытаемся понять, как мозг генерирует субстанцию *С*. Сомневаюсь, что нам удастся найти инженерное решение, позволяющее создать магическое ментальное поле. Но с инженерной точки зрения отлично оправдает себя создание предсказывающей поведение машины, у которой в вычислительное ядро встроен именно такой упрощенный конструкт.

Теорию схемы внимания, наряду с другими механистическими теориями, иногда обвиняют в обесценивании сознания или даже в отрицании его как иллюзии (а мне есть что сказать об иллюзионизме, дождитесь следующих глав). Но поскольку в теории схемы внимания сознание удается понять механистически, то и его практические применения тоже вполне возможно понять. Теория отнюдь не отрицает сознание, она помещает его в самый центр наших способностей. Как я уже говорил в нескольких предыдущих главах, сознание — это древняя, предельно упрощенная внутренняя модель, разработанная в ходе эволюции с двумя основными целями. Первая его функция, вероятно, — моделирование себя самого, чтобы наблюдать, прогнозировать и управлять собственным вниманием. Вторая — ускорение социального познания, позволяющего нам моделировать состояния внимания других и таким образом предсказывать их поведение.

Мысль моя не ограничивается тем, что сознание можно понять научно или что его в конечном итоге получится создать инженерным путем, она также состоит в том, что это инструмент невероятной силы и практической значимости.

Глава

6

Мастер Йода и Дарт Вейдер: как найти в мозге сознание?

Первый вопрос, который задают о мозге и сознании, — “где?”: в какой части мозга расположено сознание?

Знание о том, где оно располагается, конечно же, не объясняет, что это такое и как оно работает. Предположим, вы не имеете ни малейшего понятия, как сохраняются данные в памяти компьютера. Специалист мог бы развинтить ваш системный блок, ткнуть куда-то внутрь и сказать: “Вот там микросхема памяти”. Эта информация не сильно изменила бы ваши знания, но вы хотя бы поняли, с чего начать. Теперь вы знаете, куда смотреть в поисках ответов. Хотелось бы надеяться, что определение места сознания в мозге принесет подобную пользу.

Чаще всего в качестве вместилища сознания предлагается кора головного мозга, та его часть, которая наиболее сильно разрослась в ходе эволюции[1]. Второе место занимает таламус[2]. Как я отмечал ранее, он тесно связан с корой, постоянно обменивается с ней данными и каждая область таламуса в основном связывается с определенной частью коры[3].

Третье предполагаемое местоположение — загадочная мозговая структура, которая называется “ограда”: тонкий слой клеток непосредственно под корой с двух сторон мозга, рядом с ушами[4]. В 1987 г., когда я был наивным желторотым студентом и только начинал работать в нейробиологической лаборатории, я спросил своего научного руководителя, можно ли мне заняться изучением ограды мозга — тем более что о ней никто толком не знал. Он не возражал, но ограда оказалась настолько маленькой и тонкой, что от нее трудно было получить надежные сигналы, так что я переключился на другие проекты. Я до сих пор питаю нежность к ограде и всегда читаю научные новости, в которых она упоминается. Она связана с корой мозга петлей обработки данных, которая не слишком отличается от таламо-кортикальной петли. Но пусть даже мы проследили все связи, функция ограды остается неизвестной. То же самое можно сказать о многих структурах мозга. Одна из прелестей нейронауки как раз и заключается в том, что бо́льшая часть мозга — неисследованная территория.

Все три упомянутые мозговые структуры так тесно связаны друг с другом, что образуют единую систему. Правильнее было бы сказать, что человеческое сознание — скорее всего, функция коры головного мозга+, где плюс означает “плюс дополнительные структуры, такие как таламус и ограда, без которых она функционировать не может”. Поскольку кора удобно разложена по поверхности мозга, ее намного легче изучать, нежели другие структуры, и поэтому на нее направлены основные экспериментальные усилия. Сейчас перед исследователями стоит вопрос, свой вклад вносят в сознание все части коры или одни области важнее других? Какие бы зоны ни играли

здесь главные роли, в процессе осознания также предположительно будут участвовать связанные с ними области таламуса, ограды и многих других мозговых структур.

Нейронаука иногда грешит тем, что называют кляксоведением: рисует красную кляксу на поверхности мозга, приписывает ей какую-то функцию и делает вид, будто получила новое научное знание. Кору традиционно подразделяют на области размером с почтовую марку, причем у каждой есть определенные свойства. Но по мере того как ученые всё лучше разбирались в работе коры, они стали понимать, что та состоит из распределенных сетей, а не изолированных зон. Сопряженные области налаживают друг с другом функциональные связи, подобно граничащим странам, которые открывают дипломатические каналы, и эти альянсы между зонами меняются в зависимости от психологического состояния человека и умственной задачи, которую он решает.

В этой главе я опишу некоторые экспериментальные попытки выяснить, какие именно корковые сети отвечают за сознание. Не буду претендовать на то, что смогу описать все или хотя бы бо́льшую часть блестящих экспериментальных работ, проведенных по этой теме. Я выбрал лишь несколько примеров, чтобы создать общее представление о глубоких, иногда скрытых понятийных затруднениях, присущих исследованиям сознания.

Нейробиолог Сабин Кастнер, как и я, работает в Принстонском университете, но этажом ниже. Она специалист мирового уровня по зрительному вниманию. Неудивительно, с учетом нашего сотрудничества, что я разработал теорию сознания

вокруг современных данных о внимании. К слову сказать, доктор Кастнер не просто мой соавтор — она моя жена. Не счесть случаев, когда разговор за ужином переходил на темы сознания и мозга. Украшение в центре нашего стола — трехмерная распечатка модели мозга нашего 11-летнего сына (созданная на основе томограммы). Когда мы с женой иногда пускаемся в восторженные рассуждения о его сложности, сын закатывает глаза и пытается вернуть разговор на интересные ему темы (фауна позднего мелового периода).

Несколько лет назад в типовую программу выступления моей жены с докладом входила развлекательная демонстрация. Сабин раздавала слушателям 3D-очки, в которых каждый глаз видит свое. Затем она показывала картинку с мастером Йодой, которая подавалась на правый глаз, и другую с Дартом Вейдером — ее подавали на левый. Зрители не могли одновременно видеть и мастера и Вейдера. Происходило то, что называется бинокулярным соревнованием. Сначала вы видите Йоду. Через пару секунд он бледнеет, потом исчезает и появляется Дарт. Еще через пару секунд меркнет Дарт и торжествует Йода. Такое вот бинокулярное проявление инь-ян Силы.

Бинокулярное соревнование — один из первых инструментов, при помощи которых ученые пытались определить расположение сознания в мозге. Оно до сих пор остается среди самых распространенных[5]. Предположим, вы поместили человека в сканер для магнитно-резонансной томографии (МРТ), чтобы измерить активность мозга, и нашли участочек зрительной коры, нейроны которого избирательно реагируют на Йоду. Может быть, они обрабатывают

только зеленые изображения. Если показывать одного лишь мастера, нейроны активируются. Если показывать Дарта Вейдера — умолкают. Теперь предположим, что вы показываете двойное, соревновательное изображение — Йоду одному глазу, Дарта второму — и измеряете активность этих избирательных к зеленому нейронов.

Поскольку мастер Йода все время присутствует в поле зрения правого глаза, нейроны могут быть активны постоянно. Если это так, они, вероятно, связаны не с сознательным восприятием зрительной информации, а с первичной обработкой изображения таким, какое оно предстает глазу. Но предположим, что эти нейроны активируются тогда, когда образ Йоды входит в сознание зрителя, и умолкают, как только тот сообщает, что мастер поблек и в сознание вошел Дарт. В этом случае активность нейронов соответствует осознанию Йоды. Возможно, эти нейроны, активируясь и обрабатывая образ мастера, *вызывают* у человека его осознание. Очевидно, что тогда должен быть и другой набор нейронов, который, активируясь, вызывает осознание Дарта Вейдера. На первый взгляд эксперимент кажется совершенно логичным.

Такой метод впервые применили в 1990-х гг. при исследованиях мозга обезьян[6], и его все еще используют на человеческом мозге, чтобы определить зоны зрительной системы, которые соответствуют сознанию (а возможно, и порождают его)[7]. Изначально ученые предположили, что корковые зоны внизу иерархии, через которые информация попадает в систему обработки (например, зона V1), будут реагировать на изображения от обоих глаз одинаково устойчиво, никак не указывая на свое отношение к

сознанию. Они же выглядят так, будто обрабатывают данные на слишком примитивном уровне, которого бы не хватило для осознания. А вот зоны на верхушке иерархии могли бы реагировать, переключаясь между двумя изображениями, отражая то, как переключается между ними сознание. Эти зоны обрабатывают зрительную информацию более сложным, глубоким, целостным образом, а следовательно, именно в них должно было бы рождаться ее сознательное восприятие. Вот и рубежу следует проходить где-то между этими двумя полюсами. А значит, по мере того как зрительная информация поднимается по иерархическим уровням обработки, делаясь сложнее и глубже, она должна преодолевать порог и входить в сознание.

Но этого простого результата никто не увидел. На самом деле эффекты оказались размазаны по всей зрительной системе[8]. Не только вся кора, но даже более ранние стадии обработки — участки таламуса, соединяющие глаз с корой, — показывают некоторые корреляции с осознанным восприятием[9]. Их активность регулируется между правым и левым глазом по мере того, как человек осознает изображение, поданное на каждый из них. Чем на более высокий уровень корковой иерархии мы смотрим, тем сильнее нейроны коррелируют с осознанным зрительным восприятием, но все-таки кажется, что простой границы между осознанными и неосознанными зонами не существует.

Так и хочется сделать вывод, что сознание постепенно возникает из всей зрительной системы, накапливаясь в ходе того, как информация поднимается выше по уровням обработки. При таком раскладе не существует конкретной точки в мозге,

которая порождает сознание. Это свойство всей системы. Может ли такое объяснение оказаться верным или в эксперименте на бинокулярное соревнование, несмотря на его изящество, скрываются концептуальные демоны?

Предположим, наша задача — разобраться в том, как люди видят цвет. В каком-то смысле он подобен сознанию. Мы не считаем, что цвет — это вещество или что он имеет вес, но он существует и может быть приписан предметам в окружающем нас мире. Далее я опишу точку зрения на цвет, которая безусловно неверна, но логика размышлений похожа на ту, которую люди иногда принимают, говоря о сознании.

Представьте, что вы смотрите на яблоко. Традиционный нейробиологический вопрос будет звучать так: “Куда в мозге попадает информация о яблоке, чтобы создать осознаваемый зрительный опыт?” Зададим параллельный вопрос: куда в мозге попадает информация о яблоке, чтобы создать красноту?

Зрительная система обрабатывает форму предметов. V1 — зона коры в самом низу иерархии обработки — занимается простыми элементами формы, такими как угол наклона фрагмента линии на границе формы. Выше по иерархии, в зонах коры высших уровней, сходятся более сложные интегральные аспекты формы. Давайте сформулируем теорию цвета (я знаю, что она неверна, наберитесь терпения), по которой тот возникает в результате обработки формы мозгом. Согласно этой теории, когда в систему начинает входить зрительная информация, у нее нет цвета и она кодируется в черно-белой схеме. Теперь предположим, что где-то в ходе обработки, по мере того как данные о форме

обрабатываются более глубоко, начинает возникать цвет. К тому времени, как информация достигает высших уровней зрительной иерархии, он уже полностью из этих данных сформирован. В сей странной теории цвет — сущность, создаваемая мозгом, побочный продукт обработки информации о форме. Назовем ее эмерджентной теорией цвета.

Если мы согласимся с ней, следующий наш шаг — разобраться, какие части коры головного мозга создают цвет. Чтобы это понять, проведем эксперимент с бинокулярным соревнованием, измеряя активность мозга испытуемого в сканере для МРТ. Мы будем следовать той же логике, что и в описанном ранее эксперименте на сознание. В качестве конкурирующих изображений возьмем красный квадрат и красный круг. Квадрат подается на правый глаз, круг — на левый. Рождается соревнование: участник эксперимента на несколько секунд видит красный квадрат, затем квадрат блекнет и возникнет красный круг и т.д., то один, то другой. Свойство красноты сначала придается квадрату, затем кругу — по очереди. В какой-то момент информация о квадрате достигает ключевых областей мозга, в которых она создает опыт красноты. В следующий миг должна дойти до нужных зон мозга информация о круге, которая создаст красноту. Где эти зоны?

Предположим, мы найдем зону коры, которая активна, когда испытуемый сообщает, что видит красный квадрат, но умолкает, когда он говорит, что видит красный круг. Видимо, эта зона и порождает цвет. Ее нейроны обрабатывают информацию о квадратности, и, когда они возбуждены, их активность также должна вызывать красноту, возникающую из квадратности. Предположительно, другой набор нейронов — активных, когда испытуемый сообщает,

что видит красный круг, — должен выполнять такую же “краснопорождающую” функцию для круга.

Проводя этот эксперимент, мы обнаруживаем, что характерные сигналы соревнования между кругом и квадратом распределены по всей зрительной коре, от входных зон до высших уровней обработки информации. Данные о форме, где бы они ни оказались в системе обработки, создают цвет. Из этого мы делаем вывод, что цвет — распределенное комплексное свойство, которое нельзя изолировать в какой-либо части мозга. Он возникает из всей системы в целом.

Я понимаю всю абсурдность эмерджентной теории цвета и слабость предложенного эксперимента с бинокулярным соревнованием. Я описываю их здесь не для того, чтобы посмеяться над прежними подходами к изучению сознания, а с целью продемонстрировать, насколько легко пойти по ошибочному пути в рассуждениях. В конце концов, до того как сложилось научное понимание цвета, эмерджентная теория могла бы показаться правдоподобной. Выдвигались и более причудливые предположения. Сегодня мы лучше понимаем цвет, но на удочку аналогичной эмерджентной теории сознания попасться все еще легко. Однако в подобных рассуждениях возникают некоторые логические проблемы.

Первая заключается в том, что цвет — не эмерджентное свойство. В каком-то смысле цветовое зрение “возникает” из нейронных сетей. То же самое можно сказать о любом свойстве, моделируемом в мозге. Но, для того чтобы наш испытуемый сообщил, что квадрат красный, что-то в его голове должно было обработать данные именно о красном цвете. Можно

создать зрительную систему, которая будет сколь угодно сложной, интегральной и всеобъемлющей, умеющей обрабатывать информацию о форме сверху вниз, снизу вверх и во все стороны, но я ручаюсь, что если вы не включите туда часть, которая занимается обработкой именно цвета, то чувствительной к нему зрительная система не окажется. Цвет не будет иметь для нее значения. Предположение, что, если достаточно интенсивно обрабатывать квадратность (или сделать ее невероятно сложной, или перевести обработку в особые высокоуровневые зоны мозга), то она начинает порождать цвет — в корне неверно. Раз цвет — призрачная субстанция, создаваемая при обработке формы, раз нет специфически цветовой информации, которая обрабатывается подобно любому другому свойству стимула, — то как сможет человек вербально сообщать о нем? В конце концов, чтобы делать утверждения, мозг должен располагать информацией, на которой они основываются. С целью понять, как люди видят красный квадрат, не стоит надевать шоры и смотреть только на те области мозга, которые обрабатывают данные о форме квадрата, надеясь среди них найти порождающие также и цвет. Нет, нам нужно искать те проводящие пути и сети, в которых обрабатывается информация именно о цвете.

Аналогично этому, чтобы разобраться в сознательном восприятии зрительной информации, нет смысла присматриваться только к тем областям мозга, в которых она обрабатывается, в надежде найти те самые, в которых создается сущность сознания. Нужно искать сети, в которых складывается информация именно о сознании.

Второй изъян нашего выдуманного эксперимента состоит в том, что оба стимула обладают цветом и к тому же он *одинаковый*! Независимо от того, что

воспринимает наш испытуемый — красный квадрат или красный круг, — все равно идет обработка красного. Все части зрительной системы, занятые обработкой цвета, будут реагировать на оба стимула одинаково: они не станут активироваться или замолкать, когда восприятие переключится с квадрата на круг. Если и существуют зоны коры, специализирующиеся на цветовой информации, — а похоже, в коре довольно много очагов обработки цвета[10], — то наш эксперимент никак не поможет их обнаружить. Подобное происходит, когда мы изучаем сознание. Какое бы изображение ни осознавал испытуемый — поданное на правый глаз или на левый, — в обоих случаях он осознает изображение. Все области мозга, обрабатывающие информацию о наличии сознания, будут одинаково активны в обоих случаях. Они не должны активироваться или замолкать в зависимости от того, какое изображение в данный момент побеждает в бинокулярном соревновании. Такой эксперимент начисто упустит области мозга, занятые моделированием сущности сознания.

Третий недостаток эксперимента — неспособность понять источник бинокулярного соревнования. Оно вызвано соперничеством[11]. Сигналы от обоих глаз начинают взаимодействовать практически сразу после входа в мозг. Это взаимодействие разрастается по мере прохода информации в глубь зрительной системы. Если изображения, приходящие от двух глаз, заметно различаются (как в случаях мастера Йоды с Дартом Вейдером или круга с квадратом), то система не в состоянии свести их воедино и они начинают глушить друг друга.

Сигналы от правого и левого глаз — словно равные по силам борцы. То один берет верх, то другой. Ни один не может одержать окончательную победу. Сигнал, который временно побеждает, захватывает корковую систему и топит соперника. Когда обнаруживаются конкурентные эффекты в таламусе, или V1, или на любой другой ступени зрительной иерархии, нет данных, которые бы указывали, что сознание само по себе присутствует, моделируется или обрабатывается в конкретной области мозга. Данные показывают только, что два зрительных сигнала, по одному от каждого глаза, пребывают в постоянном соревновании на всех уровнях зрительной системы[12]. То изображение, которое побеждает во всесистемном состязании, в итоге приобщается к сознанию, но нам отнюдь не становится понятнее, как мозговая система выстраивает сознание или как оно связано со зрительной информацией.

Использование бинокулярного соревнования для изучения сознания — хорошая иллюстрация того, как эксперимент может на первый взгляд произвести впечатление изящного, но стоит вдуматься в его глубинные предпосылки, и он превращается в научную трясину. Логические ошибки могут казаться очевидными при обсуждении цветового зрения, но те же ошибки вкрадываются с удивительной легкостью и даже воспринимаются интуитивно удобными, когда мы размышляем о сознании. Тем не менее бинокулярное соревнование — удивительный феномен, который важно изучать на уровне нейронов и вовлеченных в него зон мозга. Многие ученые, в числе которых и моя жена, продолжают заниматься этим явлением из чисто научного интереса и ради новых знаний о зрительном соревновании. Но оказывается, оно не подходит в качестве инструмента

поиска сознания в мозге. Нам нужен другой эксперимент.

Я хочу предложить намного менее хитроумный подход, нежели эксперимент с бинокулярным соревнованием. Он неточен, грубоват, зато практичен. Начнем с поиска в мозге цветообрабатывающих сетей, а затем распространим тот же ход мышления на поиск сетей, обрабатывающих сознание.

Если бы я хотел найти очаги обработки цветовых данных в коре, я бы поместил испытуемого в сканер для МРТ и показал бы ему ряд черно-белых картинок. Иногда среди них встречались бы цветные. И если бы некая область коры подсвечивалась сильнее при появлении цветного изображения, я бы понимал, что нужно получше изучить именно ее.

Эксперимент обнадеживающе прост. Общее правило таково, что чем меньше в эксперименте рабочих деталей, тем больше его шансы на успех. Но в то же время предлагаемый эксперимент несовершенен. Прежде всего, черный и белый — тоже цвета, мозг обрабатывает их так же, как любые другие. В эксперименте сравниваются изображения меньшего и большего цветового диапазона в надежде на то, что разница между ними достаточна, чтобы подсветить зоны, обрабатывающие цвет. Второй недостаток в том, что при исследовании может высветиться множество посторонних областей мозга, никак не связанных с цветом. Например, некоторая область участвует в управлении бдительностью и подсвечивается, когда цветное изображение неожиданно появляется после целой минуты скучных черно-белых картинок. Но, несмотря на все его

несовершенства, я бы принял этот метод хотя бы как первый взнос в копилку ответов.

Такую технику уже применяли, и она отлично работает для определения основных сетей обработки цвета[13]. Сам по себе этот эксперимент — только начало, но на те же области мозга указывают и другие методы. Например, если обследовать людей, у которых эти области повреждены в результате инсульта, то обнаружится утрата цветовосприятия[14]. Такие люди видят мир в оттенках серого и даже не могут вспомнить или по-настоящему понять, что такое цвет.

Один и тот же вопрос исследуется во множестве несложных экспериментов с разных точек зрения, и результаты становятся все более достоверными. В науке, по моему опыту, лучше не зацикливаться на создании одного-единственного совершенного эксперимента с учетом всех-всех параметров и с идеальными результатами, перевязанными ленточкой. Научный успех обычно приходит с серией приближений и постепенно растущей уверенностью.

Теперь применим тот же несовершенный, но практичный подход к изучению сознания.

Предположим, я посажу испытуемого перед монитором. Он сфокусируется на точке в центре экрана. Внезапно, совсем ненадолго, появится изображение лица. Оно будет неярким и возникнет лишь на двадцатую долю секунды. Затем экран покроется случайным набором цветных квадратов. Такая зрительная маскировка, как ее называют, прервет обработку мозгом предшествующей информации: испытуемому станет труднее увидеть лицо. Оно будет балансировать на грани осознания. Каждые несколько секунд в тех же затрудненных

условиях будет показываться новое изображение. В ряде случаев испытуемый подтвердит — да, в этот раз я увидел лицо (назовем эту ситуацию условием А), но иногда он скажет — нет, лица я не видел (назовем это условием Б).

И вот мы получили пусть простой, грубый и несовершенный, но все же способ найти сознание в мозге. В условии А испытуемый зрительно осознаёт лицо и цветную маску, которая покрывает экран после лица. В условии Б он видит только маску, не лицо. В каком-то смысле у него больше осознания зрительной информации (или он осознает больше различных изображений) в условии А. Если мы найдем область мозга, которая более активна в условии А, то, возможно, именно она и участвует в сознательном восприятии зрительной информации. В любом случае это будет полезный первый шаг. Такой эксперимент может указать и на посторонние области мозга, не связанные с сознательным восприятием зрительной информации, но у нас будет хотя бы начальный вклад в ответ на наш вопрос, и тогда мы сможем перейти к другим экспериментам, например к исследованию пациентов с инсультами, у которых повреждены эти области мозга.

Вероятно, эксперимент с лицом никогда не ставили именно в описанном мною виде, но исследования такого типа, изучающие сознательное восприятие зрительной информации, проводились не единожды[15]. В условии А испытуемый осознает изображение. В условии Б предъявляется такое же или похожее изображение, но испытуемый его не осознает.

Вы, возможно, предположите, что в условии А (когда лицо достигает сознания) зрительная система

должна обрабатывать информацию о лице, а в условии Б (когда этого не случается) зрительная система не доходит до его обработки. Но происходит, как правило, не это. Области мозга, которые, как нам известно, обрабатывают данные о форме, цвете, текстуре и других видимых свойствах, непрерывно активны в ответ на изображение независимо от того, осознает ли его человек. При этом есть заметные различия между осознаваемым и неосознаваемым условиями. Активность зрительной системы обычно намного более устойчива и менее изменчива в осознаваемых условиях[16]. Эта разница активностей соответствует общему представлению о том, что, когда зрительные сигналы усиливаются и стабилизируются вниманием, у них больше шансов достичь сознания. Самая большая разница наблюдается в ряде зон мозга в теменной и лобной долях[17]. Эти зоны в целом активируются в ответ на лицо в условии А, когда испытуемый сообщает, что осознает его, но намного менее активны (если активны вообще) в условии Б, когда человек его не осознает.

Подобные результаты были многократно получены в различных вариантах экспериментов, поэтому я говорю об этих данных с определенной долей уверенности. В одном из моих собственных недавних экспериментов мы получили практически те же самые результаты[18]. Когда вам что-то показывают и вы это осознаете, активируются теменно-лобные сети. Когда вам что-то показывают и ваша зрительная система это обрабатывает, но вы это не осознаете, теменно-лобные сети не реагируют. Простейшее объяснение заключается в том, что теменно-лобные сети или какие-то подсистемы внутри них, возможно, порождают сознание.

На роль главного очага сознания часто предлагают одну из составляющих теменно-лобного комплекса — префронтальную кору. Это самая передняя часть коры головного мозга, находящаяся непосредственно за лбом. Она до сих пор не слишком хорошо изучена, но уже имеющиеся о ней сведения позволяют предположить, что там происходит обработка данных на очень высоком уровне. Отличительное свойство префронтальной коры — ее универсальность[19], она задействована практически в любой задаче, которая стоит перед человеком. Если вы точки считаете — нейроны префронтальной коры отслеживают их число[20]. Если вы наблюдаете за размещением точек на экране, те же самые нейроны перенастраиваются на местоположение[21]. В знаменитом эксперименте Эрла Миллера из Массачусетского технологического института обезьян обучали считать неоднозначно идентифицируемых по изображениям животных либо кошками, либо собаками[22]. Нейроны префронтальной коры вскоре превратились в детекторы кошек и собак: одни реагировали, когда обезьяна давала понять, что видела кошку, другие активировались, когда обезьяна видела собаку. Сама случайность задачи (обследуемые животные никогда не видели настоящих кошек и собак) показывает, что префронтальные нейроны могут обрабатывать любые свойства окружающего мира в зависимости от поставленной задачи.

Нейробиологи иногда называют префронтальную кору, особенно одну ее большую подструктуру — дорсолатеральную префронтальную кору, — отделом “рабочей памяти” в мозге[23]. Данные от других зон мозга могут загружаться в префронтальную область, где информация временно хранится и обрабатывается для тех нужд, что актуальны прямо сейчас.

С учетом этих ее свойств префронтальная кора, похоже, идеально соответствует тому, чтобы оказаться театром сознания[24]. Она собирает информацию по распределенным связям со всей остальной коры, словно паук в центре паутины.

Однако, сколь бы заманчиво ни выглядело это решение проблемы сознания, меня оно не убеждает. Остается слишком много вопросов. Например, почему переживание сознания не пропадает при повреждениях префронтальной коры? Да, у пациентов с такими повреждениями возникают трудности с планированием будущего и переключением с одной задачи на другую, но обычно они сохраняют общее ощущение сознания[25]. По крайней мере, его утрата не является типичным отличительным свойством префронтальных повреждений.

Но у меня есть даже более общий вопрос к префронтальной гипотезе: почему информация, достигая этой области мозга, вообще создает ощущение сознания? Пока нейробиологи не найдут обиталище сознания — мозговой центр, в котором информация сводится воедино и рождает сознание, — наука не сдвинется с мертвой точки.

Это та же проблема, что я очертил раньше. Куда идет информация о форме, чтобы создать цвет? Никуда! Чтобы понять цвет, нужно смотреть туда, где обрабатывается информация о цвете. Куда идет зрительная информация, чтобы создать осознанный опыт? Никуда! Чтобы понять сознание, нужно искать в мозге систему, которая обрабатывает *информацию о сознании* — его свойствах и эффектах. Информация о сознании, которая обрабатывается в одной системе, связана со зрительной информацией о разглядываемом вами предмете, которая

обрабатывается в другой системе. Два комплекта данных, объединенные во временном союзе мозговых сетей, строят крупномасштабную внутреннюю модель, которая, по сути, говорит: “Это мой сознательный опыт, а это — яблоко, и прямо сейчас одно соединяется с другим”.

Я ничего не имею против того, чтобы считать префронтальную кору участницей этого процесса. Возможно, как и подозревают ученые, она служит дискуссионной площадкой нашего ума — активно собирает наши цели, мысли и наблюдения, помогает организовать наше поведение. Это предположение соответствует имеющимся данным и обладает определенной убедительной силой. Но мне кажется, что собрание сигналов со всех концов мозга должно где-то в себе содержать и информацию о сознании, иначе мы бы не могли делать никаких утверждений об этом свойстве. Мы уже знаем, какие сети в мозге обрабатывают зрительную информацию о яблоке. Зрительную систему тщательно изучали почти целый век. Но где те сети, которые обрабатывают информацию о сознании?

Коллеги-нейробиологи знают, как расплывчато я выражаю свои мысли при обсуждении нейронных сетей, протянувшихся от теменной доли к лобной. Некоторые хорошо изучены[26], например: дорсальная и вентральная сети внимания, сеть выделения перцептивной значимости стимулов (насыщенности), управляющая сеть, сеть модели психического состояния, сеть пассивного режима работы мозга (default mode network). Их намного больше, причем некоторые выполняют узкоспециализированные функции. Одна из сетей связана с управлением движениями глаз и, вероятно, перекрывается по

составу с той, что управляет вниманием[27]. Другая связана с управлением протягивания рук[28]. Благодаря третьей сети руки принимают нужное положение, чтобы брать предметы[29]. До того как переключиться на изучение сознания, я в своей лаборатории изучал ту теменно-лобную сеть, которая обрабатывает данные о пространстве в непосредственной близости к телу (“окололичностном” пространстве) и участвует в координации защитных действий, если что-нибудь придвигается слишком близко к нему[30]. Еще одна сеть предположительно участвует в счете и математическом мышлении[31]. Разнообразие сетей невероятно. Вдобавок к этому изобилию у многих сетей нет четких границ, и они участвуют в решении множества различных пересекающихся задач. Где в этом хаосе сетей и областей может обрабатываться информация о сознании?

Сейчас я склонен предполагать, что при моделировании, относящемся к построению концепции сознания, особенно много активности наблюдается в той области коры, которая носит название “височно-теменной узел” (TPJ — temporo-parietal junction)[32]. В мозге два таких узла, по одному с каждой стороны, на поверхности коры, соответствующей зоне непосредственно над ушами. Узел на правой стороне мозга может быть крупнее или лучше развит, чем на левой, но природа этой асимметрии пока неясна. В каждом полушарии мозга этот узел подразделяется на зоны с разными свойствами, причем особенно меня интересует самая верхняя — та, что лежит в основном в теменной, а не в височной доле. В нашей лаборатории мы зовем эту зону дорсальным височно-теменным узлом, или TPJd (см. рис. 6.1). Также иногда ее называют нижней

теменной долей. Дорсальный височно-теменной узел расположен на “перекрестке” управляющей сети, вентральной сети внимания, сети выделения перцептивной значимости и сети модели психического состояния[33]. Это один из информационных центров мозга с самым большим количеством связей — он мог бы даже посоревноваться по этому параметру с префронтальной корой. Если он и впрямь вносит вклад в процессы, порождающие сознание, то, скорее всего, работает не в одиночку — сложное моделирование обычно зависит от распределенных сетей, а не от четко определенных зон. Скорее всего, этот узел служит и множеству других целей помимо указанной выше. В дальнейшем я буду ссылаться на височно-теменной узел в целом, а не на его дорсальную область, так как у нас недостаточно данных для исключения влияния остальных частей узла. Но мы подозреваем, что именно TPJd может играть особенно важную роль в порождении сознания.

Давайте я проясню, в чем именно состоит мое предположение. Вероятно, височно-теменной узел сам по себе ничего не осознает и не порождает сознательный опыт. Это не гомункул — спрятанный в голове маленький человечек, — а только часть той сети, которая строит концепцию, модель, а также информирует мозг, что такое сознание. Без этой информации мы не могли бы приписывать самим себе сознание, утверждать, что обладаем им, или хотя бы понимать, о чем речь, если нас о нем спрашивают.

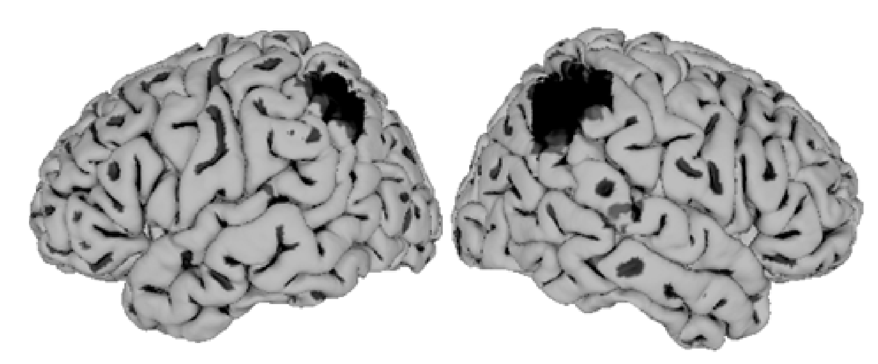


Рис. 6.1. Верхняя, или дорсальная, область височно-теменного узла (TPJd) на левой и правой сторонах коры головного мозга человека, определенная объединением двух методов: сначала применялась техника анализа независимых компонентов, затем определялась схема связей с остальными областями коры (воспроизводится в соответствии с изображением на с. 9435 статьи K. M. Igelström, T. W. Webb, and M. S. A. Graziano, “Neural Processes in the Human Temporoparietal Cortex Separated by Localized Independent Component Analysis,” *Journal of Neuroscience* 35 (2015): 9432–45)

У меня есть несколько причин уделить особое внимание височно-теменному узлу. Первая заключается в том, что эксперименты последних 30 лет показали его вовлечение в обработку возможных психических состояний других людей[34]. Когда вы размышляете о намерениях, убеждениях и эмоциях окружающих, ваш височно-теменной узел активен более других в сети построения модели психического состояния. Если мы ищем области мозга, которые способны раскрыть для нас само понятие сознающего разума, естественно искать их в височно-теменном узле. Когда люди думают о сознании в ходе экспериментов моей лаборатории — отвечают на вопросы о своем собственном сознании (“Да, я осознаю эту картинку”) или приписывают осознанные состояния другим (“Да, я думаю, что Кевин осознает тот предмет, который лежит рядом”), — мы видим реакцию височно-теменного узла[35]. Он так активен, будто участвует в построении сознания.

Вторая причина, по которой я подозреваю, что височно-теменной узел может вносить вклад в сознание, — это его взаимоотношения с вниманием.

Вспомним, что в теории схемы внимания мозг создает информацию о сознании по конкретной причине. Эта информация служит полезной схемой внимания, моделью, описывающей некоторые его свойства. Любая сеть в коре головного мозга, которая участвует в управлении вниманием, должна иметь доступ к схеме внимания. Следовательно, нужно смотреть на связанные с ним сети — дорсальную и вентральную сети внимания, сеть выделения перцептивной значимости, управляющую сеть. Считается, что эти сети — двигатели внимания, они способны влиять на иные области коры, чтобы усиливать одни сигналы за счет других[36]. Естественно предположить, что именно там будет строиться схема внимания. В ряде экспериментов, включая проведенные в моей собственной лаборатории, мозговая активность, связанная с вниманием, и мозговая активность, связанная с социальным познанием, демонстрируют характерную область пересечения в височно-теменном узле, особенно в его верхней части[37].

Сомнения, касающиеся вклада височно-теменного узла в сознание, основывались на том, что при осознании человеком зрительного образа, например лица, не было свидетельств поступления данных о лице в височно-теменной узел или вообще куда-либо в теменную долю. Ни черты лица, ни его выражение, ни составляющие его тонкие видимые детали, как считается, не обрабатываются в теменной доле. В теории 1990-х гг., теперь ставшей классической, психологи Дэвид Милнер и Мел Гудейл[38] предположили, что теменная кора *не связана* с сознательным восприятием зрительной информации, тогда как другие области зрительной системы, особенно в височной доле, *участвуют* в ее осознании.

Другие ученые отвергали предположительную роль и теменной, и височной долей как источников сознательного восприятия зрительной информации, утверждая, что сознание соединено с нижними уровнями обработки информации, где та впервые попадает в кору в зоне V1[39]. Прямо сейчас, глядя на экран компьютера, я осознаю изгибы и углы букв, которые набираю. Поднимая взгляд, я вижу листву платана за окном. Мне доступны ее яркие цвета, я различаю тончайшие детали и текстуру листьев. Зона V1 извлекает из видимого именно такие подробности. Более того, при повреждениях этой зоны люди утрачивают осознанное зрительное восприятие[40]. С учетом этих наблюдений можно допустить, что сознание зарождается не на вершине корковой иерархии, где представлены мышление и познание, а в самом низу, где обрабатываются сырые подробности.

Подобные взгляды будто бы противоречат моему предположению, что в сознании участвует височно-теменной узел. Теория схемы внимания, однако, перестраивает имеющиеся предпосылки и позволяет всем точкам зрения состыковаться. Предположим, вы зрительно осознаете изображение лица. Зона V1 помогает выстроить данные о визуальных подробностях и цветах, насыщенности и яркости. Височная доля — данные о чертах лица для его распознавания. Височно-теменной узел — данные о сознании. Все эти компоненты необходимы, чтобы вы могли заявить, что осознаете лицо. Милнер и Гудейл были правы в том, что *содержание* сознательного восприятия зрительной информации (в данном случае подробности и черты конкретного лица), как правило, не обнаруживается в теменных зонах. Но без височно-

теменного узла не могло возникнуть и само понятие сознания, с которым бы соединилось содержание.

Традиционно зрительную систему представляют себе по типу сборочного конвейера: пакеты данных входят в глаз, затем следуют от одной стадии обработки к другой, поднимаясь по ступеням иерархии и получая все более глубокую и полную проработку. Когда продукт достигает высшего уровня, его создание закончено, и предыдущие стадии, чья работа завершена, больше не нужны. Теперь высший уровень корковой иерархии может взять полностью собранные данные и превратить их в действия или речь — или, может быть, в сознание.

Возможно, здесь больше подойдет аналогия со стопкой чертежей для постройки дома. На каждом их листе здание описывается по-разному и с разной степенью детализации. На одном листе представлены контуры помещений. На другом — схема разводки электричества. На третьем — устройство канализации, а на четвертом — встроенная мебель. Есть даже лист, где описываются предлагаемые отделочные материалы. Чертежи образуют естественную иерархию, поскольку ни один лист нельзя начертить, не разобравшись в деталях предыдущих. Информация в каком-то смысле течет от листа к листу, проходя все более глубокую обработку. Но в то же время, пролистывая чертежи, вы не отбрасываете предшествующие и не летите на всех парах к итоговому генеральному плану, в котором на одном-единственном листе будет собрана вся необходимая информация. На каждом листе есть что-то полезное, и если потерять хоть один, то дом построить не получится. Итоговое действие — постройка дома — возникает из каждого листа проекта, а не только из последнего, итогового, на

котором собрана максимально обработанная информация.

Подобным же образом обладает иерархией и зрительная кора — скорее, даже множеством взаимосвязанных иерархий[41]. Но каждый следующий уровень не замещает предыдущие. Нет конечной стадии, на которую после долгой обработки в итоге попадает все важное, чтобы достигнуть сознания или запустить действие. Напротив, каждая стадия иерархии делает свои выводы, опуская одни данные, обогащая другие, добавляя новые слои к успешному анализу видимого мира. Такое накопление информации — от мельчайших деталей, отраженных на низших уровнях, к глубоким концептуальным данным, представленным на высших, — способно влиять на действия и речь. Результат выходит не только из высшего уровня иерархии, но от всех уровней одновременно.

Когда кто-то говорит: “На этой фотографии — мой друг”, такое утверждение основывается на информации, поступившей в области обработки лиц на высших уровнях зрительной иерархии, в височной доле. Когда кто-то говорит: “У моего друга пятнышко на носу”, то подобное замечание, скорее, основывается на информации от зрительных областей низшего уровня, где обрабатываются мелкие визуальные детали. Когда кто-то говорит: “У моего друга темные волосы”, это основано на информации от обрабатывающих цвет областей в специфических цветовых очагах зрительной коры. Когда кто-то говорит: “У моего друга скрипучий голос”, то подобное суждение может основываться на информации, извлеченной из памяти и воспроизведенной в одной из областей слуховой коры.

Когда кто-то говорит: “Завтра днем я верну ему должок в 20 баксов”, это может основываться на информации, частично выстроенной в связанной с планированием префронтальной коре.

А если человек говорит: “У меня есть *субъективное переживание* всего этого содержания. У меня есть загадочное субъективное «я», сознание”, то такое утверждение базируется на информации, выстраиваемой в еще одной специализированной корковой сети, которая может проходить через височно-теменной узел.

Я вовсе не имею в виду, что кора головного мозга — это набор отдельных модулей, каждый из которых ограничен информацией из одной сферы и изолирован от всех остальных. Нет, зоны коры, хоть и специализируются на разных типах информации, взаимодействуют друг с другом и влияют друг на друга, образуя огромную единую резонирующую сеть. Троньте одну из ниточек этой сети — добавьте новую зрительную информацию от зоны V1 или новую мысль от префронтальной коры, и нить вызовет отклик по всей сети, вызывая переменчивую активность. Пакеты информации усиливаются вниманием или тонут в фоновом шуме. Зоны коры постоянно образуют союзы, открывая временные каналы сообщения, объединяя свою информацию в более крупные структуры, а затем разъединяются, чтобы вступить в новые союзы.

При таком подходе нам нет нужды выбирать, с чем сознание связано сильнее всего — с низшими ступенями иерархии, с высшими или же с чем-то в середине. Мы можем осознавать и визуальные подробности, которые обрабатываются в зоне V1, и абстрактные размышления, которые обрабатываются

в префронтальной коре. Причина в том, что понятие сознания определяется комплексом информации, построенным в специализированной сети, а эта сеть может образовывать функциональные союзы со множеством других зон коры.

Представьте себе, что на Землю прилетел разумный инопланетянин. Предположите, что в его мозге нет концепции сознания, которое объединяет всю остальную полученную информацию. У него нет аналога человеческого височно-теменного узла. Вероятно, у нашего инопланетянина все же есть какие-то модели себя — он ведь может наблюдать за происходящими внутри него процессами и управлять своим поведением, но его эволюция шла другим путем, и его модели себя не содержат понятия субъективного сознания.

Мы спросим инопланетянина: “Видишь яблоко?”

Он ответит: “Да, вижу”.

Мы: “У тебя есть сознательный опыт яблока?”

Инопланетянин: “Я обрабатываю его зрительно”.

Мы: “Но есть ли у тебя еще кое-что, вот это дополнительное *je ne sais quoi*[10], сущность опыта, ассоциируемого с яблоком, — или ты просто обрабатываешь информацию? Ты что-нибудь чувствуешь внутренне, когда обрабатываешь информацию о яблоке?”

Инопланетянин: “Я не знаю, что это за дополнительное *je ne sais quoi*. Вот яблоко, я обрабатываю его зрительно. Что мне еще делать? Почему должна быть какая-то сущность сверх этого? Из чего она состоит и откуда берется?”

Мы: “Ага, но ты пользуешься словом «я», значит, у тебя есть представление о себе!”

Инопланетянин: “Ну да. Этот разговор происходит между двумя действующими лицами, которых можно обозначить как «ты» и «я». У меня есть вполне достаточный набор данных о моем физическом теле, моем прошлом и поведенческих характеристиках”.

Мы: “То есть ты все-таки осознаешь себя?”

Инопланетянин: “Я знаю о себе. Что такое осознавать?”

Мы: “Это когда вдобавок к обычной обработке информации о себе ты переживаешь ее еще и субъективно. У тебя есть нечто дополнительное, такое *je ne sais quoi*… а ладно, забей”.

Я хочу сказать, что разум — это информация. Я люблю говорить, что человеческий разум — суперскрученная из миллионов нитей информационная скульптура, постоянно меняющаяся, прекрасная в своей сложности. Если у разума нет информации о сознании, он не может знать, что из себя представляет это свойство, не может приписать его себе или другим. Мы, люди, придаем такое значение сознанию только потому, что в нашем мозге есть подсистема, которая выстраивает информацию о нем. И не исключено, что наука приближается к пониманию того, где именно расположена эта подсистема и как она связана с остальными системами.

Старый добрый способ[11] определять, где в мозге локализованы различные функции, — изучать последствия повреждений мозга. Например, зрительную кору выявили, когда поняли, что повреждения задних частей мозга обезьян в зоне,

которая теперь известна как V1, вызывают слепоту[42]. Речевые центры в человеческом мозге были открыты, когда выяснилось, что повреждения зоны лобной доли, теперь известной как речевая зона Брока, отнимают у людей способность говорить[43]. Участки мозга, зрительно отслеживающие движение, были определены, когда стало понятно, что повреждения определенной части зрительной коры вызывают слепоту к движению[44]. (Известен случай пациентки с таким нарушением, из-за которого она не могла налить воду в стакан: все, что пациентка видела, это сосульку, свисающую из кувшина, а затем — лужу на столе.) Повреждения же цветообрабатывающих зон коры приводят к утрате цветовосприятия[45].

Может ли повреждение определенного участка коры привести к утрате сознания?

Речь не о том, чтобы человек уснул. Циклы сна и пробуждения управляются намного более древними структурами в стволе мозга[46]. Речь о состоянии, в котором человек технически бодрствует, информация входит в корковую систему и даже, возможно, влияет на поведение, но не происходит никакого субъективного осознания этой информации.

Речь не идет и о философском зомби[47]. Философы придумали довольно специфический его концепт — это не пожиратель мозгов, который ковыляет с вытянутыми вперед руками, а человек, который нормально выглядит и ведет себя, даже нормально разговаривает, но не имеет внутренних субъективных переживаний. Ученые обсуждают возможность существования зомби. (Условный инопланетянин, которого я описывал выше, — именно он.) Теория схемы внимания оставляет место для подобной возможности: зомби можно было бы создать

искусственно или они могли бы появиться на другой планете в ходе эволюции. Но, если эта теория верна, мы никогда не сможем, удалив часть мозга обычного человека, превратить его в философского зомби, который ходит и разговаривает. Нам не удастся удалить сознание, не лишив человека возможности нормально функционировать в обычной жизни. Для этого нормального функционирования человеческому мозгу и необходим такой конструкт, как сознание.

Если бы вам хирургически удалили схему внимания, вы бы столкнулись по меньшей мере с тремя нарушениями. Во-первых, вы бы не смогли управлять собственным вниманием. А если вы не в состоянии управлять своим вниманием, то вы абсолютно недееспособны, не можете направлять свои действия последовательным и устойчивым образом на конкретные предметы или цели. Вы беспомощнее, чем зомби из фильмов, которые хоть и хромают, но удерживают свое внимание на том, чтобы убивать людей и есть их мозги. Во-вторых, вы бы не были способны строить социальные модели сознания других людей. Без возможности приписывать сознание другим развалилось бы ваше социальное познание. В-третьих, без необходимого комплекса информации вы бы не могли понимать вопросы, касающиеся сознания, или делать о нем осмысленные утверждения. Традиционный философский зомби может слиться с толпой, но у “зомби схемы внимания” есть серьезные нарушения. Если его кольнуть булавкой, он, может быть, и отреагирует, но в остальном вряд ли станет что-то делать.

В клиническом мире повреждений мозга известен особенно тяжелый и разрушительный синдром — одностороннее пространственное игнорирование

(неглект). Его изучают уже почти сто лет[48]. Повреждение одного из полушарий мозга приводит к утрате осознания предметов и событий с другой стороны тела. Те ключевые элементы, которые, повреждаясь, вызывают этот синдром, есть в обоих полушариях мозга, но нарушения справа обычно вызывают более обширную и постоянную потерю осознания[49].

Игнорирование — это не слепота. При отсутствии видимости предметы исчезают из поля вашего зрения, но вы не теряете знания о них. Скажем, все мы слепы относительно предметов, находящихся за спиной, но у вас, скорее всего, есть неплохое представление о том, что прямо сейчас находится сзади вас. При игнорировании же предметы, которые попадают в “плохую” половину пространства, попросту исчезают из сознания. Не распознается прикосновение к левой стороне тела. Звук, исходящий слева, или приписывается источнику справа, или вообще не замечается. В особенно тяжелых случаях пациент вообще не осознает, что существует левая половина пространства.

Не будет ничего удивительного в том, что пациент с игнорированием побреется только с одной стороны, попытается одеть половину тела, съест еду с половины тарелки. Если тарелку после этого повернуть, он не поймет, откуда взялась лишняя еда. Когда мы попросим его нарисовать картинку, он нарисует правую ее сторону и не заметит, что половины картинки не хватает. При предложении нарисовать часы он, скорее всего, нарисует полный круг (возможно потому, что сохранил моторную память о круговом движении руки). Но все числа, от 1 до 12, он

упихнет в правую его половину и будет думать, что всё сделал правильно[50].

Когда я учился в аспирантуре, мне довелось пообщаться с одним таким пациентом, зашедшим в нашу лабораторию. Вообще-то мы не занимались клиническими исследованиями, но самый простой тест знали. Мы взяли лист бумаги и нарисовали в разных его местах множество коротких горизонтальных линий. Потом мы положили лист перед пациентом и попросили его зачеркнуть все линии, которые он видел.

Помню, как он засмеялся и сказал: “Опять то же самое!” Он всмотрелся в лист и зачеркнул все линии в его правой половине. Когда мы спросили, *уверен* ли он, что закончил, он снова взглянул на лист и сказал: “Точно, я несколько пропустил”. И зачеркнул еще пару линий ближе к центру листа, оставив всю левую половину нетронутой. Мы повернули лист так, что лево оказалось справа, и он поразился, как много линий пропустил. Этот тест изумлял его каждый раз. Пациенты с игнорированием знают, что что-то не так, но не могут понять, что именно.

В одном из самых жутких экспериментов с такими пациентами проводили тест на зрительное воображение[51]. Участников просили представить, что они стоят на северной стороне знакомой им площади, и назвать все здания, которые те смогут вспомнить. Они тут же перечисляли здания, которые оказались бы справа от них. Затем пациентов просили представить, что они стоят на южной стороне той же площади. Тогда они называли “противоположный” набор зданий. Пациенты знали, что отвечают неверно, но не могли понять, откуда берется ошибка. Они даже не могли представить себе левую сторону той площади.

Но информация с игнорируемой левой стороны не исчезает бесследно. В одном особенно печально-выразительном эксперименте[52] пациентка с игнорированием разглядывала изображение дома, из окна на левой стороне которого вырывались языки пламени. Когда ее попросили описать картинку, пациентка описала обычный дом. Она не осознавала огонь, но сказала, что дом ей не понравился. Что-то было не так, даже если она не могла понять, что именно. Этот эксперимент — всего лишь один из многих, где демонстрируется, что информация с игнорируемой стороны попадает в мозг, обрабатывается и может косвенным образом влиять на поведение. Если уколоть пациента булавкой с игнорируемой стороны, он поморщится и поймет, что случилось что-то неприятное, но не осознает конкретной боли. Если с левой стороны бросить в него мячом, пациент может, сам не понимая почему, пригнуться. При игнорировании половина пространства не стирается, но не осознается, и на нее не получается направить внимание.

В контексте теории схемы внимания простейшее объяснение игнорирования состоит в том, что механизм внимания, в том числе и схема внимания, может быть разделен по крайней мере на два поля — правое и левое. С правой стороны пространства пациент способен оставаться более-менее жизнеспособным в том, что имеет отношение к вниманию, поведению, социальному познанию и утверждению о субъективном сознании. С левой стороны пациент становится зомби схемы внимания.

Одностороннее игнорирование обычно возникает после инсультов, повреждающих обширную область коры, поэтому трудно определить, какая именно

корковая зона отвечает за его основные симптомы. Обнаружение областей, поврежденных у многих пациентов со слегка различающимися симптомами, сузило область поиска эпицентра игнорирования. Игнорирование различной степени тяжести возникает при повреждениях самых разных зон, но наиболее серьезные и длительные нарушения возникают при повреждении височно-теменного узла, особенно той его части, что расположена в теменной доле[53]. Если повредить эту ключевую зону, сознание распадается.

Одностороннее игнорирование, пусть и разрушительное для пациентов, действует все же только на половину пространства. У них остается часть сознания. Они могут взаимодействовать с близкими и “ощущать жизнь” посредством другой половины пространства. Во многих случаях со временем синдром ослабевает, и по мере реорганизации[12] сетей мозга у пациентов восстанавливаются если не все, то хотя бы часть утерянных возможностей.

Но у некоторых пациентов после инсульта повреждения мозга оказываются настолько обширными, что бедняги полностью утрачивают сознание. К простейшим реакциям многие из них остаются способны. Если кольнуть таких пациентов булавкой или подуть на сетчатку, они отреагируют. Их мозг не умер — бóльшая его часть осталась в целости и работоспособной сохранности. Но они не дают никаких реакций на других уровнях. Они будто больше не приходят в сознание. Это самые душераздирающие случаи — когда утрачивается суть личности.

Мозг пациентов, находящихся в этом “вегетативном состоянии”, часто сканируют, чтобы

определить масштаб повреждений. Они могут быть обширными и различаются от пациента к пациенту. Но становится видна закономерность: есть ключевой набор зон мозга, при повреждениях которых чаще всего наблюдается утрата сознания: теменно-лобные сети[54]. Если удалить эти сети с обеих сторон мозга — вы удалите сознание. Зомби схемы внимания — не философская фантазия “а что, если”, не мысленный эксперимент. Это медицинская трагедия из реальной жизни.

Глава

7

Трудная проблема и другие точки зрения на сознание

Может показаться, что сознание — это предмет ученых обсуждений и что в мире существует слишком много разных, не связанных друг с другом точек зрения. Такое впечатление подкрепляется и распространенным представлением о том, что сознание — неразрешимая загадка. Я настроен оптимистичнее. Я полагаю, что мы приближаемся к пониманию механизма сознания и, возможно, даже уже разобрались в его основных принципах. Оптимизм мой основывается в том числе на том, что в этом массиве разнообразных теорий и точек зрения многие различаются не так сильно, как кажется на первый взгляд. Если копнуть поглубже, обнаруживаются многообещающие совпадения.

В этой главе я объясню, как теория схемы внимания соотносится с семью другими широко известными научными взглядами на сознание. С некоторыми из них я буду спорить, а с некоторыми найду пересечения. Конечно, в ограниченном объеме одной главы я не в состоянии привести полное изложение всех альтернативных взглядов. Моя задача — оценить не отдельный подход, а, скорее, соотношение его и теории схемы внимания.

Трудная проблема и метапроблема

Упоминавшийся выше термин “трудная проблема” предложил философ Дэвид Чалмерс. И этим он задал направление научных дебатов о сознании на десятилетия вперед[1]. Сознание — трудная проблема, поскольку это личный опыт, который невозможно подтвердить извне. В силу самой природы субъективного опыта на него нельзя надавить и измерить противодействие, нельзя положить его на весы и узнать, сколько он весит, или нагреть его и найти температуру сгорания. Он непроницаем для науки. Трудная проблема, на самом деле, это эвфемизм для *невозможной* проблемы, к решению которой наука приблизиться не может.

Позже Чалмерс также предложил термин “метапроблема”, отсылающий к вопросу, почему мы вообще считаем, что у нас есть трудная проблема[2]. Может быть, ее и нет. Может быть, внутри нас не существует принципиально необъяснимой нефизической субстанции. Может быть, наша задача как ученых — объяснить, почему люди в принципе склонны верить в трудную проблему.

Теория схемы внимания — пример второго подхода. В сущности, эта теория объясняет, почему биологическая машина ложно убеждена в наличии трудной проблемы. Когда машина обращается к своей схеме внимания — упрощенному, карикатурному изложению внутренних процессов, — она получает информацию, что в ней существует скрытая, призрачная сущность сознания.

В ядре теории кроется особая мощь — знание, основанное на моделях и обладающее гораздо

большей силой, нежели поверхностное. Чтобы стало понятнее, приведу пример.

Вот девочка играет, воображая себя щенком. Ребенок тявкает и передвигается на четвереньках, говоря: “Я щенок!” Чтобы девочка могла сделать такое заявление, ее мозг должен был сформировать ключевое утверждение “я щенок”, а также у него должны были быть сведения о том, что щенки тявкают и ходят на четырех лапах. Но эти данные существуют в более широком контексте. Мозг ребенка содержит обширную сеть информации, в том числе “на самом деле я не щенок”, “я это выдумала, это понарошку”, “я маленькая девочка” и т.д. Какая-то часть этой информации присутствует на когнитивном и лингвистическом уровнях. Намного больше ее находится на более глубоком сенсорном, или перцептивном, уровне. Схема тела девочки строится автоматически, значительно ниже высшего уровня мышления, — и описывает она физическое строение тела человека, а не щенка. Девочка видит свои человеческие руки, зрительная информация подтверждает ее человеческую сущность. Она помнит, как ела ложкой хлопья на завтрак, сидела на уроке, читала книжку — все это человеческие действия. Утверждение “я щенок” — поверхностное, оно не совпадает с ее глубинными внутренними моделями.

Но предположим, что у меня есть фантастический способ изменить информацию в ее мозге. Я изменю схему тела девочки, которая станет соответствовать телу щенка. Я изменю информацию в ее зрительной системе и памяти, чтобы они соответствовали утверждению “я щенок”. Я уберу когнитивные данные, которые сообщают: “Я это придумала, это понарошку”. Я подменю информацию, гласящую: “Я уверена, что это неправда”, ее противоположностью.

Как тогда девочка поймет, что она не щенок? Ее мозг — заложник содержащейся в нем информации. Выходит тавтология — он знает только то, что знает. Девочка уже не будет считать свою щенячью личину игрушечной или ненастоящей. Она станет воспринимать ее буквально, как истину. У нее не появится поводов думать как-то иначе.

Вы можете попробовать переубедить ее. К примеру, вы скажете: “Но ты же понимаешь человеческий язык и умеешь говорить. Щенки так не могут. Тебе не кажется, что это повод задуматься — нет ли здесь путаницы в том, кто же ты?”

Предположим, что девочка наша оказалась интеллектуально одаренной и понимает логику наших рассуждений. Новая информация попадет на поверхностный, когнитивный уровень. Она войдет в конфликт с более глубокими внутренними моделями. Девочка будет интуитивно верить в одну истину о себе, а интеллектуально принимать во внимание другую.

Так и я, пишущий эту книгу, возможно, смогу убедить вас, что ваше сознание коренится в схеме внимания. Интеллектуальное рассуждение означает, что раз вы заявляете, что осознаете, значит, у вас есть информация об этом. Но интуитивно вы верите в иную правду о себе. Когда вы полагаетесь на интроспекцию и обращаетесь к той самой схеме внимания, она рассказывает вам другую историю. Информация в ней говорит вам: не-е-е-т, ваше сознание — это не данные, не механизм, не нейроны, а эфирная субстанция, неотъемлемо присущее вам свойство, живущее внутри вас. Предположим, я выступлю со своими рассуждениями убедительно. Тогда вас начнут раздирать противоречия:

поверхностное, интеллектуальное знание будет вести к одному пониманию себя, а глубинные внутренние модели — к другому. Вы никогда не сможете устранить этот разрыв. С помощью поверхностного интеллектуального знания или нескольких часов размышлений вам не удастся уничтожить схему внимания, которая формировалась миллионы лет эволюции и поэтому встроена в вас очень глубоко.

Еще один пример глубоко встроенной в нас несовершенной модели — то, как мы видим белый цвет. Согласно модели, которую строит зрительная система, белый — это яркость в отсутствие загрязняющих цветов. Данная модель развивалась миллионы лет, причем так же она устроена у многих видов животных. В какой-то момент, а именно в 1671 г., одно особенно мозговитое животное по имени Исаак Ньютон разобралось, что сия внутренняя модель — упрощение[3]. Белый свет — смесь всех цветов, а мозг отражает его в упрощенном виде.

Трудную проблему белого цвета можно было бы сформулировать так: “Что за особый физический процесс очищает белый свет от всех загрязняющих его?” Вот соответствующая ей метапроблема: “Почему мы вообще думаем, что существует трудная проблема? Почему мы считаем белый цвет очищенным?” Теперь мы знаем решение метапроблемы: мозг строит простую, практичную, но несовершенную модель. И отсюда мы понимаем, что трудную проблему решать не надо.

И тем не менее, хотя любой образованный человек теперь знает, что белый — смесь всех цветов, это знание не меняет моделей, встроенных в зрительную систему. Мы все же видим белый чистым, а не смесью всех цветов. И никто не возмущается противоречием.

Мы привыкли к тому, что интеллектуальное, когнитивное знание не согласуется с тем, что сообщают нам глубинные, врожденные модели мозга. Можно сказать, что наука — постепенный процесс, в ходе которого когнитивные части нашего мозга обнаруживают неточности в глубинных, эволюционно встроенных моделях мира.

Философ Франсуа Каммерер задал о теории схемы внимания мудрый вопрос[4]. Предположим, теория верна. Мозг строит схему, которая как бы имитирует внимание. Она обрисовывает общие свойства высшего уровня, в частности нашу способность сосредоточиваться и глубоко обрабатывать информацию. Но в то же время при этой обрисовке схема не учитывает физических или механистических свойств внимания. Она не утверждает, что у внимания нет физической сущности, — просто замалчивает эту тему. В схеме нет данных о мелких деталях вроде нейронов и синапсов. Если наши представления о сознании основаны на этой внутренней модели, почему мы тогда так уверены, что сознание — это эфирная сущность? Откуда мы берем представление о том, что сознание ничего не весит и не имеет материального воплощения, если об этом нет данных в соответствующей внутренней модели?

Разгадка, мне кажется, в том, что люди в целом не склонны к такому представлению. Мы не считаем сознание чем-то физически несуществующим, а понимаем его как нечто, к чему материальные свойства *не имеют отношения*. Но ведь это совсем другое представление.

Чтобы лучше понять меня, представьте себе, что кто-то тронул вас за плечо. Прикосновение активирует на коже рецепторы, передающие информацию в мозг.

В конечном итоге тот строит определенного вида внутреннюю модель — тактильную модель, пакет информации, описывающий это конкретное прикосновение. Модель содержит данные о месте прикосновения, его силе в начальный момент, о давлении, продолжительности и, возможно, даже о мягкой фактуре кончиков пальцев. Это богатая сенсорная репрезентация. Но она не содержит информации о вкусе. Прикосновение к плечу не обладает, например, соленостью. Я не говорю, что прикосновение пресно и его стоило бы подсолить. Нет, оно попросту лежит *вне* измерения вкуса. Оно расположено в другом информационном пространстве. Теперь, когда я обратил ваше внимание на такую возможность, вы сможете обдумать ее на поверхностном, когнитивном уровне, но у вас не получится поменять глубинную внутреннюю модель. Тактильное восприятие — врожденный процесс, недоступный когнитивным изменениям. Вам не удастся придать прикосновению вкус.

У меня практически нет сомнений, что если бы можно было внедрить электроды в мозг здорового человека вам удалось прочитать информацию, зашифрованную в его тактильной системе, то в модели восприятия *не* содержалось бы информации: “А вкус, кстати, отсутствует”. Выраженного отрицания даже и не требуется. Вкусовые качества просто обойдены молчанием. Наше интуитивное представление о прикосновениях не связано с тем, что вкусовые ощущения в них отключены, мы воспринимаем прикосновения как нечто, к чему вкус попросту не имеет никакого отношения.

Полагаю, что схема внимания действует так же. Она содержит богатый, но ограниченный набор информации. Она отражает общие свойства

внимания, но не физические, не механические. Основываясь на этой внутренней модели, мы интуитивно представляем себе внутренний мысленный опыт, который может завладевать информацией и подталкивать нас к действию, как это делает внимание, но этот опыт не имеет никакой определенной связи с материальным воплощением. Его нельзя физически схватить, он не гладкий, не шершавый, не волнистый, не тяжелый, не легкий, не вонючий, не зеленый, не заостренный. Он нигде не расположен в этих физических измерениях — точно так же как прикосновение не существует в измерении солености.

И все же в этом представлении схема внимания описывает по меньшей мере одно физическое свойство. Она описывает внимание как нечто, имеющее физическое расположение где-то внутри нас. Имея внутреннюю модель такого рода, мы должны представлять себе мысленную сущность, которая накладывается на материальный мир, в том смысле, что мы можем указать на определенное место и сказать: “Она существует примерно здесь”. Это что-то вроде призрака, обитающего в физическом пространстве несмотря на то, что у него нет никаких других физических свойств. Согласно этой теории, “призрак в машине”, сознательная энергия внутри нас, — это представление, возникающее напрямую от схемы внимания, в которой данные о внимании неполны.

И вот мы снова возвращаемся к трудной проблеме и метапроблеме. Трудная проблема возникает вследствие допущений, сделанных на основе той глубинной модели — схемы внимания. Теория схемы внимания — метаответ, который объясняет, почему

вообще люди верят в существование трудной проблемы.

Иллюзии и метафоры

Является ли сознание иллюзией?

Иллюзионизм — относительно новый и активно развивающийся теоретический подход к сознанию[5]. Основной его посыл состоит в том, что действительного, реального сознания у нас нет. Само по себе переживание опыта, субъективная сущность — отсутствует. Просто мы *считаем*, что сознаем: это иллюзия, созданная мозгом. Может быть, она дает какое-то функциональное преимущество (по одной из теорий, придает жизни огонька), а возможно, она вовсе не нужна для выживания и возникает как случайное следствие обработки информации в мозге[6]. Когда-нибудь мы поймем, какой в ее основе лежит механизм, и тогда разберемся, есть ли в ней функциональный смысл. А пока ученым нет нужды объяснять, как возникает несуществующее сознание, точно так же как нет нужды объяснять, почему Земля плоская или почему Солнце крутится вокруг Земли. В науке появляется все больше сторонников этого взгляда, но такую теорию людям трудно принять.

Теория схемы внимания — тоже своего рода иллюзионизм. Она утверждает, что наиболее трудного для понимания свойства сознания — его эфирной, метафизической природы — в реальности не существует. Мы считаем, что обладаем этим свойством, только потому, что нас дезинформирует несовершенная внутренняя модель.

Но, как показывает мой опыт, назвать сознание иллюзией — это поставить на теории крест. Вас, быть может, поймет горстка философов, но остальной мир отмахнется от вашей теории, сочтя ее оторванной от действительности ерундой: “Как это сознание может быть иллюзией, когда у меня в голове столько всего творится?”

Слово “иллюзия” до того легко неправильно понять, что оно может стать непреодолимой преградой в обсуждениях сознания. Я сейчас рассмотрю три ловушки применения этого слова к сознанию. Но при этом я не хочу развенчивать идею, лежащую в основе иллюзионистского подхода, — в целом она мне кажется верной.

Помню, на пляже играли в песке маленькие мальчик и девочка, лет пяти. Мальчик очень серьезно сказал: “Нам нельзя долго быть на солнце, а то у нас вырастут клешни”.

Девочка изумилась. “Правда, что ли?” — переспросила она, уставившись на собеседника.

Мальчонка торжественно кивнул: “Правда”. Он поднял руки и изобразил движения клешней. “Мне мама сказала!”

Прелесть этой истории в том, что мальчик очевидным образом неправильно понял общеупотребительную метафору. Мама, скорее всего, сказала сыну, что тот станет похожим на рака — имея в виду, что он покраснеет, обгорев на солнце. А воображение мальчика нарисовало не красный цвет, а клешни.

Метафоры следуют строгим подразумеваемым правилам[7]. Обычно основа (рак) сравнивается с объектом метафоризации (обгоревший на солнце

человек). Значение имеет лишь одно свойство. У рака их много: клешни, экзоскелет, фасеточные глаза на стебельках, но от человека, к которому обращена метафора, ожидается, что он поймет, какое то самое ключевое свойство предназначено для переноса. Мы все интуитивно пользуемся метафорами именно так.

Когда ученый говорит: “Сознание — это иллюзия”, то, как мне думается, большинство людей неявно трактуют это утверждение как метафору. Основа метафоры — зрительная иллюзия — обладает многими возможными свойствами. Предмет может казаться больше, чем есть на самом деле, или более наклоненным, или более удаленным. Неподвижный предмет — восприниматься движущимся. Выпуклая поверхность — вогнутой. Но, когда слово “иллюзия” используется в контексте метафоры, оно как будто может значить только одно. Люди выделяют единственное ключевое свойство: они приравнивают иллюзию к миражу. Когда вы видите мираж, вам кажется, что нечто присутствует, хотя на самом деле его нет. Вы не ошибаетесь в размерах или деталях, вы ошибаетесь в самом существовании предмета.

Вот пример. Представьте, что друг жалуется вам: “Профессионализм моего начальника — чистая иллюзия!” Он не имеет в виду, что начальник профессионал, но не совсем такой, как вы могли бы ожидать (а может быть, он даже лучше, чем вы думали). Нет, друг хочет сказать вам, что его начальник на самом деле вообще не компетентен. Называть нечто иллюзией в контексте метафоры — значит отрицать само его существование.

Если вы заявите, что сознание — иллюзия, большинство людей воспримет это совсем не так, что “сознание по сути напоминает зрительную иллюзию:

настоящая обработка информации в наших головах слегка отличается от картины, которая представляется нам на основе интроспекции, поскольку мы делаем свои выводы на основе несовершенной внутренней модели — все это напоминает зрительную иллюзию”. Нет, это поймут в том смысле, что за иллюзией ничего не стоит. Там нет даже “там”[13]. Сознания не существует, никого нет дома. Такое представление покажется большинству людей абсурдным.

Как я уже писал в предыдущих главах, в теории схемы внимания сознание — никакой не мираж. Это упрощенное, несовершенное, но все же отражение реально существующего явления. Мозг действительно захватывает информацию и глубоко ее обрабатывает. Когда мы сообщаем, что переживаем осознанный опыт, мы даем слегка схематизированное изложение этой буквальной истины. На самом деле, там есть “там”. Сознание напоминает иллюзию в техническом смысле, поэтому теория схемы внимания является иллюзионистской технически. Но сознание не является иллюзией в том смысле, в котором эту метафору поймет большинство.

Давайте рассмотрим иллюзии поближе и разберемся, в чем их отличие от нормального восприятия. Ваша зрительная система строит внутренние модели, упрощенные репрезентации предметов окружающего мира. Она делает это постоянно, автоматически, всё то время, пока у вас открыты глаза. Нормальный акт зрения сам по себе — не иллюзия. Иллюзия — это особый случай; он имеет место, когда система делает ошибку. Маленький предмет кажется большим, прямой — наклонным. Что-то глючит. Большинство исследователей зрения трактуют термин “иллюзия” как отклонение, искажение нормы.

Предполагаемая схема внимания — стандартная внутренняя модель, которая строится непрерывно и автоматически. Она дает вам упрощенные данные о состоянии вашего внимания. Скорее всего, она может дать и сбой: тогда глюк внутренней модели породит иллюзию. Видеть сознание в кукле — иллюзия. Но я бы не стал называть иллюзией нормально работающую внутреннюю модель. Любая такая модель — упрощенная версия действительности, а действительность эта настолько сложна и имеет такое детальное строение, что обрабатывать ее у мозга нет ни причин, ни ресурсов. Если любую упрощенную внутреннюю модель считать иллюзией, то избыточны термины “зрительная” и “иллюзия”: иллюзией тогда будет все зримое. Если таково новое определение этого неоднозначного слова, то все, что мы видим, слышим, думаем и чувствуем, — иллюзия. Вероятно, философы, которые называют сознание иллюзией, имеют в виду как раз такое всеобъемлющее определение: сознание — часть несовершенного понимания действительности мозгом. Я могу понять логику их рассуждений, но это всеобъемлющее определение все же воспринимаю скептически. Если применять слово ко всему, оно потеряет свое значение.

Главная опасность, возникающая при назывании сознания иллюзией, — возможность неверной трактовки такого вывода: в нем можно усмотреть логическую закольцовку. Для большинства людей иллюзия по определению — разновидность сознательного опыта. Если сознание — иллюзия, то кто переживает иллюзию? Получается, что сознание объясняется через сознание.

Подобная критика обескураживает, поскольку основывается на неверном понимании. Для

философов-иллюзионистов в мозге нет ничего, что испытывало бы иллюзию сознания. Это мозг утверждает, что обладает сознанием, — на основе неполной информации.

Выходит увлекательно: когда я заговариваю с философами, они выглядят как бы опешившими: “*Конечно же,* нет в мозге ничего, что испытывало бы субъективный опыт иллюзии сознания. Мы не это имеем в виду! Иначе мы бы попали в порочный круг”. А потом я заговариваю с кем угодно еще, из любой другой сферы жизни, и они с тем же ошарашенным видом отвечают: “*Конечно же,* иллюзия подразумевает, что ее должно переживать нечто сознающее. Что еще означает это слово? Зачем пользоваться им, если вы имеете в виду что-то другое? Почему, например, не назвать сознание летучей мышью и не создать свой собственный язык?”

Поэтому я не называю теорию схемы внимания иллюзионистской, хотя с точки зрения иллюзионистов она именно такова. Наши разногласия касаются терминологии, а не содержания. Может быть, сознание больше напоминает карикатуру, ведь оно искажает нечто реальное. Не уверен, правда, что можно что-то существенно прояснить, назвав сознание карикатурой вместо иллюзии. Хлесткое название или слоган для сознания, скорее всего, никогда не смогут дать ответ, которым мы удовлетворимся.

Некий невероятно сложный, тонко устроенный механизм обладает основанным на моделях знанием о себе и мире. Одна такая внутренняя модель — схема внимания — дает нам интуитивные представления о сознании. Большинство свойств, которые мы приписываем сознанию, существуют в

действительности и поддаются измерению в мозге — в виде высших уровней коркового внимания. Некоторые из свойств, приписываемых сознанию (его метафизическая, эфирная природа), — результат неполных или упрощенных данных во внутренней модели.

И что, в соответствии с этими свойствами сознание представляет собой иллюзию? Да, можно назвать его так, если дать термину аккуратное определение. Думаю, что мои друзья-иллюзионисты в философии подходят к этому с умом и точны в своих определениях. Я совсем не хочу на них нападать, напротив — надеюсь на то, что смогу показать, как значительно наши взгляды могут объединиться в рамках единой теоретической модели.

Фантомные конечности

Если вас когда-нибудь, не дай бог, угораздит потерять конечность, возможно, вы станете ощущать фантом — вам будет казаться, что отсутствующая конечность все еще на месте. После ампутации кратковременные фантомные ощущения возникают у 90% пациентов, а у некоторых они могут тянуться годами[8]. Сохраняются ощущения поворота в суставе, прикосновений, боли, холода и тепла. Глазами вы видите, что конечности больше нет, но ощущения остаются от каждого ее сантиметра. Словно свой собственный призрак, конечность незримо тянется из тела. Великий британский адмирал лорд Нельсон, потерявший руку в битве при Санта-Крус-де-Тенерифе, во всеуслышание заявлял, что доказал существование загробной жизни: раз у его руки есть дух, значит, он есть и у всего остального тела[9].

Фантомные конечности — не просто медицинские диковинки. Это крайне неприятный опыт. Представьте себе, что ваша конечность психологически совершенно реальна, но вы не можете почесать ее, размять суставы или облегчить боль. Переживания фантомных конечностей бывают мучительны, изнурительны — не говоря уже о том, насколько это странно и сбивает с толку[10]. К примеру, иногда фантомная конечность становится “телескопической”, и пациент сообщает, что фантомная кисть руки торчит у него прямо из плеча или что фантомная нога укоротилась и больше не касается земли.

Общепринятое объяснение феномена фантомных конечностей состоит в том, что, даже если сама конечность потеряна, сохраняется ее внутренняя модель в мозге[11]. Обширный комплекс информации, описывающий конечность, еще удерживается в проводящих путях мозга. Этот феномен показывает, какой властью может обладать внутренняя модель. Если бы не зрение, постоянно удостоверяющее отсутствие конечности, вы бы никак не могли оценить истинность своих ощущений и предполагали бы, что у вас всё на месте.

Прямая противоположность этому явлению возникает при клиническом синдроме под названием соматопарафрения[12]. Пациенты, у которых после инсульта повреждены определенные части теменной доли, могут утрачивать внутренние модели конечностей. Конечность на месте, пациент ее видит, но она будто ему не принадлежит. Невролог Оливер Сакс, у которого замечательно получается описывать клинические случаи в предельно человечной манере,

рассказывал о мужчине, потерявшем представление о своей ноге.

> — Полегче! — остановил его я. — Успокойтесь! Я на вашем месте не стал бы так ее бить.
>
> — Почему это? — воинственно и раздраженно поинтересовался он.
>
> — Потому что это ваша нога, — ответил я. — Разве вы не узнаете собственную ногу?[14][13]

Еще более впечатляющая демонстрация внутренней модели тела — исследования с резиновой рукой. Данные по иллюзии резиновой руки впервые опубликовали в 1998 г. Мэтью Ботвиник и Джонатан Коэн[14]. Я пережил на своей шкуре более позднюю версию эксперимента, пройдя его в лаборатории Хенрика Эрссона в Стокгольме. Я сидел на придвинутом к столу стуле, засунув руку в дырку в обувной коробке. На коробке лежала резиновая рука телесного цвета, она была меньше моей и смотрелась не очень-то реалистично. Я с трудом удерживался, чтобы не захихикать.

Экспериментатор надел кольцо на указательный палец моей руки, что была в коробке. Кольцо соединялось с таким же кольцом на указательном пальце резиновой руки при помощи короткого пластикового стерженька, просунутого в дырочку в крышке коробки. Каждый раз, когда я поднимал собственный указательный палец в коробке, в силу простого механического соединения поднимала указательный палец и резиновая рука на крышке.

Мне хватило пяти или шести движений пальцем, чтобы иллюзия накрыла меня с головой. Резиновая рука вдруг стала моей собственной. Меня шокировало, насколько глубоко было это ощущение. Моя схема тела вобрала в себя кусок резины, который (как я знал умом) не имел ко мне никакого отношения. Даже не знаю, как описать, до чего же жутко мне стало. Схема

тела — это не зрительная информация о руке. Это не интеллектуальное знание. И не медицинское знание о том, какие части тела с какими соединяются. Это не история, которую я рассказываю сам себе. Это внутренняя модель, которая порождается автоматически, она намного ниже когнитивных уровней, вне зоны действия сознательной воли. Эта модель дает информацию “когнитивному” мозгу, и мы оказываемся в заложниках у данных, которые она поставляет. Она сообщает, что резиновая рука — моя: р-р-раз, и я на интуитивном уровне оказываюсь совершенно в этом уверен, хотя парадоксальным образом знаю, что это не так.

Схема тела содержит информацию, которая нужна мозгу, чтобы помогать управлять движениями. Она сообщает, какие объекты принадлежат вашему телу. Эта схема описывает форму каждой его части, структуру крепления конечностей, их размеры и размах, возможности движения подопытной руки. Чего она не содержит, так это информации о механистической, детальной структуре внутри тела. Схема руки ничего не сообщает о строении костей или креплении связок, медленно- и быстросокращающихся мышечных волокнах, пролегании кровеносных сосудов или молекулах белка в клетках мышц, вызывающих их сокращение. Если вы сейчас закроете глаза и будете рассказывать о своей руке — не излагать медицинские знания о том, как она должна быть устроена, а передавать ощущения от нее в настоящий момент, — то сможете описать лишь поверхностные свойства, но не механистические подробности. Это поверхностное описание предоставляет вам схема тела.

Что такое фантомная конечность с философской точки зрения? Что такое фантом вообще? К какой категории

его отнести? Это не предмет. Не энергетическое поле. На самом деле из культи ничего не торчит. Отсутствующая конечность — вроде незримой сущности, жизненной силы, оставшейся на месте после того, как оторвали физическую плоть и кровь. Как заметил лорд Нельсон, она схожа с духом. Распространенная во многих культурах вера в духов, возможно, возникает из глубинных внутренних моделей, таких как схема тела, которые скармливают подобные концепции высшему мышлению. Но мы не можем просто взять и отмахнуться от фантомных конечностей как от ненаучных суеверий о привидениях. Они сообщают нам о чем-то важном. Еще до ампутации существует некий прикладной процесс, который обнажается с утратой конечности. Фантом — это симуляция, которая есть у всех нас, а у подвергшегося ампутации пациента симуляция сохраняется и после утраты конечности, которую эта симуляция должна представлять.

Я вижу близкое сходство между фантомными конечностями и сознанием — между схемой тела и схемой внимания. Одно — призрак в теле, другое — призрак в голове. И то и другое — симуляции. Они принадлежат к разным областям одного и того же объекта: многосоставной модели себя. Схема тела — модель физического "я" и принципов его действия, а схема внимания — модель другой части "я", взаимодействующих нейронов в черепной коробке и принципов *их* действия. Обе модели оставляют за кадром излишнюю механистическую информацию. Можно сказать, что схема внимания — специализированное расширение схемы тела.

Долгие годы я изучал схему тела — не только то, как мозг моделирует само тело, но и то, как он моделирует вокруг него безопасную подушку

пространства, которая искривляется в соответствии с движениями конечностей и головы, — словно толстый слой незримого желе[15]. Понимание того, что мозг моделирует себя упрощенным, нереалистичным, но удобным для практического применения способом, напрямую привело к теории сознания как схемы внимания.

Нейробиолог Олаф Бланке и философ Томас Метцингер предложили особенно сильное обоснование связи между схемой тела и сознанием[16]. По их мнению, телесное знание о себе — это первобытная, минимальная форма сознания, знание себя как агента, отдельного от остального мира. Очевидно, такое рассуждение отлично стыкуется с теорией схемы внимания.

Глобальное рабочее пространство и кракен сознания

В старших классах у нас была компания, в которую входили заметные, выделяющиеся из общей массы ребята. Они же были центральным узлом сплетен. По школе гуляло много дурацких слухов и обрывков информации, но, если какая-то пикантная подробность добиралась до членов этой компашки, она немедленно становилась доступна всем; вскоре уже вся школа была в курсе событий. Между прочим, именно так работает теория сознания как глобального рабочего пространства. Достигающая его в мозге избранная информация транслируется повсюду и получает возможность направлять наше поведение и речь.

Теорию глобального рабочего пространства первым предложил Бернард Баарс в 1980-х гг.[17] С тех

пор ее разрабатывали многие — в частности, Стэн Дехан[15], — чтобы привести в соответствие с современным знанием о сетях в коре головного мозга[18]. Согласно этой теории, которую я вкратце описал в четвертой главе, информация просачивается по корковой иерархии. Что-то из нее способно повысить мощность своего сигнала, обогнать другие сигналы и попасть на высшие уровни обработки — вероятно, в теменно-лобных сетях. Там информация оказывается в центральном мозговом узле сплетен — глобальном рабочем пространстве. Она достигает, по уже приведенному выше выражению Дэниела Деннета, “славы в мозге”[19]. Информация, которая попадает в глобальное рабочее пространство, попадает и в сознание.

Трудность с теорией глобального рабочего пространства состоит в том, что она не предлагает никакого объяснения тому, почему информация, попавшая в глобальное рабочее пространство, получает вместе с этим и свойство осознанного опыта. Эта теория неполна, хотя в некотором роде ее неполнота — это фича, а не баг. Можно собирать данные об анатомическом мозговом субстрате сознания, не принимая ничью сторону в философском споре. В каком-то смысле она напоминает изначальную научную теорию сознания — высказанную 2500 лет назад мысль Гиппократа о том, что за разум отвечает мозг[20]. Теория Гиппократа стала важной вехой в науке, поскольку верно выделила субстрат, хоть и не объяснила, что такое сознание и как оно возникает.

Теория схемы внимания предлагает способ завершить эту картину. Предположим, что теория глобального рабочего пространства в общем и целом

верна. Внимание может усиливать и отбирать информацию, пока та не достигнет глобального влияния в мозгу. Теория схемы внимания утверждает, что вдобавок к этому мозг конструирует схематичную модель внимания. Он строит свою собственную наивную метафизическую теорию о том, что такое глобальное рабочее пространство.

В каких отношениях тогда оказываются глобальное рабочее пространство, схема внимания и сознание? Я пытался подобрать сравнение, чтобы объяснить все сложности, и лучшая удавшаяся мне аналогия — это кракен и гигантский кальмар. Кракен — мифическое существо из скандинавской мифологии, которое впервые упомянуто в исландской саге XIII в. об Одде Стреле[21]. Он подобен огромному кальмару, наделенному сверхъестественной силой и яростью, который крушит корабли. Кракенов не существует, но существуют гигантские кальмары. Они живут в океанских глубинах, достигая длины 15 м. Их редко можно увидеть, мы мало о них осведомлены, они никогда не нападают на корабли, поскольку погибают от падения давления воды, если всплывают слишком близко к поверхности. Ноги мифа о кракене почти наверняка растут из искаженного описания гигантского кальмара.

Внимание подобно гигантскому кальмару: трудноуловимое явление, но реальное. Описывать высший уровень внимания в мозге как глобальное рабочее пространство — все равно что описывать голову гигантского кальмара как глобальный мешок с органами: удобный для ученых способ представления его анатомии. Но ничто из этого не объясняет сознание — кракена. Чтобы объяснить кракена и его значение, недостаточно препарировать кальмара и

сказать: “Вот, это кракен”. Это не кракен. Кракен — сверхъестественная, искаженная версия кальмара. Его культурное и эмоциональное влияние несравнимо больше. Полная теория кракена обязана содержать некоторое понимание кальмара, но также она должна включать в себя и понимание процесса мифологизации, который приводит к вере в кракена. Так и полное объяснение сознания не может оставаться только с вниманием и глобальным рабочим пространством. Оно должно также содержать ту наивную модель себя, схему внимания, что нашептывает нам о кракене сознания.

Мышление высшего порядка

Экземпляр моей книги, который вы читаете, скорее всего, снабжен названием на обложке, титульным листом внутри и кратким описанием на обороте. Вся эта дополнительная информация, строго говоря, не является содержанием. Это информация *о* содержании. Это метаинформация, она маркирует и обозначает. Чтобы книга вышла в свет и оказалась полезной, на ней должна быть эта метаинформация высшего порядка.

Философ Дэвид Розенталь предположил, что похожие процессы происходят и с информацией в мозге[22]. Вот я вижу яблоко. Чтобы сказать: “Я осознаю яблоко”, мне недостаточно того, что моя зрительная система обработала информацию о нем. В ходе обработки должна порождаться мысль высшего порядка, которая связывается с информацией о яблоке. Это предположение называется теорией мышления высшего порядка и укладывается в общую

систему, которую иногда именуют “мышление о мышлении”, или метапознание[23].

Ученые все еще спорят, какая именно информация высшего порядка может быть добавлена к информации о яблоке, чтобы мы стали его осознавать. Может быть, это служебные сведения[24]. Компьютерный файл содержит основную информацию — содержание файла, — а также служебную информацию высшего порядка, которая отображается как иконка на рабочем столе. Иконка обозначает файл. Это упрощенная, сжатая версия файла, лишенная подробностей. Не исключено, что наша зрительная система обрабатывает яблоко, а затем сжимает информацию в подобие иконки, которая может использоваться для когнитивного доступа и вербального сообщения. Исходя из этой точки зрения, мы утверждаем, что осознаем яблоко, поскольку для нашего высшего мышления иконка трактуется как акт сознания.

Другой вариант информации высшего порядка — оценка достоверности[25]. Допустим, вы быстрым шагом прошли мимо вазы с фруктами, среди которых было яблоко. Если вы не уверены в том, что видели, то станете утверждать, будто не осознали яблока. Если ваша оценка достоверности высока, то вы сообщите, что осознали яблоко. Вероятно, сознательное восприятие зрительной информации — это визуальная информация плюс высокая уверенность в том, что вы действительно ею обладаете.

Еще вариант — его выбрал философ Дэниел Деннет — состоит в том, что мышление высшего порядка добавляет более сложный, культурно обусловленный слой представлений[26]. Мы все являемся продуктами культуры, и, наверное, один из самых хорошо

усвоенных нами культурных мифов — тот, что у нас внутри есть душа, переживающая субъективный опыт. Если бы нас воспитывали в другой культуре, возможно, мы не имели бы понятия о сознании и я бы не писал эту книгу. Но благодаря мысли, родившейся десятки тысяч лет назад в пещере или у костра и получившей вирусное распространение среди всего человечества, мы приобрели понятие о сознании, и наше мышление высшего порядка применяет это понятие ко всему, что мы делаем. Представление о том, что сознание — это комплекс культурных мемов, также предлагала психолог Сьюзен Блэкмор[27].

Теория схемы внимания — теория мышления высшего порядка. Она очевидно принадлежит к той же категории. Однако в чем-то она отличается от иных примеров теорий мышления высшего порядка. В теории схемы внимания, когда мозг строит осознание яблока, зрительная информация о нем связывается с другими, дополнительными данными. Но эти дополнительные данные не очень хорошо вписываются в категорию “высшего порядка”. В каком-то смысле они находятся на том же уровне, что информация о яблоке. И то и другое — репрезентации реальных объектов. Мозг строит отображение формы яблока, цвета яблока, пространственного взаиморасположения вас с яблоком — и так же он строит отображение фокуса вашего внимания на яблоке. Эта схема внимания не относится к интеллектуальной, концептуальной или когнитивной сфере. Иными словами, она не относится ко второму порядку — высшему. Она не впитана с культурой, и никакое воспитание не сможет создать ее или отучить от нее. Мы не можем произвольно включать и выключать эту схему. Она развилась миллионы лет назад, задолго до появления нашего биологического

вида, задолго до речи или мышления человеческого уровня. Она сложилась на уровнях глубже языковых, хотя мы можем говорить о ней и имеем к ней некоторый когнитивный доступ. Подобно форме, цвету, видимому движению и расположению в пространстве, взаимодействие внимания между “я” и яблоком — еще один компонент обширного досье, которое мозг составляет на яблоко и окружающий его контекст. Компонент сознания в некотором смысле настолько же базовый, как и любой другой. Мы верим, что обладаем сознанием — в самом глубинном, интуитивном и иррациональном понимании, — поскольку мозг строит автоматические модели своего мира и себя самого, а также обладает частичным когнитивным доступом к этим моделям.

Конечно, я согласен, что мозг *помимо этого* выстраивает и другие мысли, высшего порядка. К примеру, мы делаем это для цвета. У нас есть эмоциональные, культурные и даже политические ассоциации с красным цветом. Эти дополнительные коннотации наслаиваются на базовую, автоматическую модель красного, которая строится в глубинах зрительной системы. Подобные культурные и личные мифы есть у нас и о сознании — откуда оно берется, зачем оно нужно, как оно соотносится с философией и духовностью и что с ним происходит, когда мы умираем. Но в теории схемы внимания под слоем культурных и концептуальных ассоциаций, которые могут быть разными у разных людей, у нас также есть врожденная схема внимания, которая более-менее одинакова у всех и дает нам общую точку отсчета сознания.

Внимание и осознание

В 1890 г. Уильям Джеймс, один из основоположников современной психологии, написал: “Каждый знает, что такое внимание. Это пристрастное, осуществляемое посредством умственной деятельности обладание в ясном и четком виде одним из нескольких, как кажется, одновременно возможных объектов или рядов мысли. Фокусировка, концентрация сознания — его суть. Это означает отказ от каких-то вещей, чтобы эффективно заниматься другими…”[16][28]

Это известное и убедительное высказывание Джеймса показывает распространенную путаницу между вниманием и сознанием. Согласно американскому психологу, у людей есть нечто, что мы называем “психический мир”, или “сознание”, а внимание относится к его сосредоточению. Тогда получается, что внимание сделано из сознания. Это центральный фокус в более широком поле сознания.

Я думаю, что представление Джеймса отражает именно то, как большинство людей понимают слово “внимание” вне научного контекста. Но это понимание довольно далеко отстоит как от современного академического использования данного слова, так и от значения его в моей книге. Разрыв между разговорным и научным значением слова “внимание” — возможно, главный источник путаницы и разногласий вокруг теории схемы внимания. Из-за этой-то двусмысленности я никогда и не любил пользоваться термином “внимание”, но он так глубоко укоренился в психологии и нейробиологии, что лучшего варианта я пока не подобрал.

В нейробиологии внимание — это процесс в мозге, в ходе которого усиливаются сигналы одной репрезентации (например, зрительного

представления яблока), ослабляются сигналы конкурентных репрезентаций, а усиленные сигналы получают соответственно большее влияние на системы в мозге. Внимание — не просто сосредоточение на одном центральном объекте: оно может быть разделено и распределено. Вам может показаться, что вы осознаете что-то находящееся за пределами внимания (т.е. уделяя внимание А, одновременно осознаете Б, В и Г на периферии), но это вряд ли может происходить на самом деле, если только не пользоваться разговорным определением внимания. По научному определению вы, скорее всего, в какой-то степени уделяете внимание всем этим предметам или быстро переключаете его между ними. В разговорной речи мы можем называть внимание разновидностью сознания, но в научном смысле их соотношение иное. Внимание — многослойный набор механизмов, метод обработки данных, а сознание — внутреннее переживание опыта, наличие которого мы утверждаем. Внимание — то, что мозг делает. Сознание — то, о наличии чего мозг заявляет.

В 1890 г. Джеймс ничего не мог знать о технической стороне обработки информации. Алан Тьюринг изложил принципы работы вычислительных машин только в 1930-х гг.[29], а Клод Шеннон сформулировал теорию информации в 1940-х гг.[30] Джеймсу никогда бы не пришло в голову рассматривать внимание как работу вычислительной машины. Он понимал его как состояние сознания, и связь между этими двумя понятиями оставалась нерушимой в умах ученых в течение еще сотни лет.

Разделение внимания и осознания впервые убедительно продемонстрировали в 1999 г. Это плоды

работы группы лондонских исследователей, среди которых были Роберт Кентридж, Чарльз Хейвуд и Ларри Вайскрантц: они изучали человека с необычными нарушениями зрительного осознания[31].

Пациент GY попал в автомобильную аварию в возрасте восьми лет и потерял практически всю первичную зрительную кору. После катастрофы он ослеп на правую сторону пространства, а также на бо́льшую часть левой. Когда взрослого GY сажали перед экраном компьютера, в разных местах которого появлялись яркие точки, пациент видел их только в небольшой области экрана левее центра. У него не было осознанного зрения за пределами этой небольшой области. Но когда его просили указать точку, которая находилась в его предполагаемой слепой зоне, то будто бы чудом он довольно точно справлялся с этим заданием. Пациент не мог осознанно видеть ее, но зрительная информация попадала в мозг и могла использоваться для того, чтобы корректно направить руку. Он также мог без сознательного восприятия зрительной информации различать базовые видимые черты, например горизонтальная перед ним линия или вертикальная. Исследуемый говорил, что ничего не видел, но просто с большой уверенностью *знал*, что перед ним. Знание представлялось ему когнитивно достоверным. Это странное явление называется ложной слепотой, оно системно встречается при повреждениях первичной зрительной коры.

Кентридж с коллегами изучал пациента GY, чтобы понять, что тот может, а чего не может делать в своей “слепой” зоне. Сначала ему показывали в этой зоне точку. Сразу после этого в том же месте показывали отрезок линии. Задачей пациента было ответить,

какую линию он видел — вертикальную или горизонтальную. В этой ситуации пациент реагировал очень быстро, поскольку, видимо, предварительно показанная точка привлекала его внимание и подготавливала к обработке линии. Но, когда экспериментаторы демонстрировали точку в одном месте, а линию — в другом, реакции исследуемого замедлялись. Вероятно, небольшая задержка с ответом возникала из-за того, что точка привлекала внимание пациента к неправильному месту и затем этому вниманию приходилось смещаться туда, где появлялась линия.

Описанный эксперимент стал поворотным для исследований сознания. Наконец удалось показать, что механизмы внимания могут продолжать работать, даже когда механизмы сознания сломаны. Очевидно, внимание — это не просто локальное сосредоточение сознания. Это иное свойство.

За последнюю четверть века множество исследований подтвердили, что осознание можно отделить от внимания — не только у людей с повреждениями мозга, но и у здоровых[32]. Людям дано обращать минимальные объемы внимания на неяркие или недолго видимые изображения (в смысле фокусировать на них ресурсы обработки информации и даже как-то реагировать), но при этом они могут утверждать, что совсем ничего не видели.

В отсутствие осознания внимание, похоже, работает не совсем обычным образом — как и следовало бы ожидать при потере части управляющего механизма. Например, один из необходимых в повседневной жизни навыков — это способность *не* фокусировать внимание на том, на чем не следует. Мир полон предметов, жаждущих нашего

внимания, которое нам иногда нужно заставить отвлечься от одного предмета (например, жужжащего над ухом комара) и сосредоточиться на другом (скажем, на книге). В подобной ситуации мы можем уделять какое-то количество внимания комару, следя за ним в фоновом режиме, но все же хотим обращать больше внимания на книгу. Такое испытание — один из самых очевидных примеров ситуаций, в которых пригодилась бы схема внимания. Чтобы справиться с задачей такого рода, мозгу нужно знать, в каком состоянии находится внимание в каждый конкретный момент, и отслеживать, когда слишком много его отвлекается от книги на комара. Также мозгу нужна рабочая модель пространственной и временно́й динамики отвлечения внимания, чтобы ему противодействовать. Теперь предположим, что в системе осознания произошел сбой. Некоторое количество внимания утекло на комара, но человек не осознает присутствия зудящего насекомого. Внимание и осознание расщепились. Если верна теория схемы внимания, то в таком случае мозг не узнаёт об утечке его на неправильный предмет. Мозговая модель внимания неполна. В таких обстоятельствах у людей, скорее всего, возникнут трудности с минимизацией внимания, уходящего на комара. Парадоксально, но неосознание насекомого приведет к тому, что на него переключится от книги еще больше внимания. Эксперименты (проведенные как в моей лаборатории, так и в других) с использованием отвлекающих стимулов, аналогичных комару, и целевых стимулов, подобных книге, подтвердили такой расклад[33].

Исходя из разнообразия и многочисленности экспериментов, может возникнуть ложное впечатление, что отделить осознание от внимания в лаборатории не составляет труда или что такое

разделение легко происходит с людьми в обычных условиях. Нет, соскрести осознание с внимания — это как отскоблить краску от стены. Обычно они крепко сцеплены. Внимание и осознание начинают расслаиваться, только если довести зрительную систему до предела ее возможностей в контролируемых лабораторных условиях, предъявляя очень короткие или размытые стимулы. Но практически невозможно подобрать зрительный стимул, который был бы настолько размытым, настолько кратковременным, настолько заслоненным другими стимулами, что люди его не осознавали бы, — и при этом отмерить его ровно такое количество, чтобы у него хватило мощности отхватить хотя бы самую чуточку внимания. Нам потребовалось несколько лет пилотных экспериментов, чтобы вбить клин между вниманием и осознанием[34] (и другие ученые рассказывали о похожем опыте).

В теории схемы внимания весь смысл осознания заключается в том, чтобы мозг получал обновляющуюся информацию о состоянии внимания. Следовательно, осознание тщательно наблюдает за вниманием — примерно как схема тела следит за положением руки. Эта парочка расходится, только если система подвергается нагрузке и с трудом балансирует на грани своих возможностей.

Интегрированная информация

Взяв яблоко, вы способны обработать его цвет, форму, запах, гладкость, звук откусывания, вкус, свою эмоциональную реакцию на него и многое другое. Каждая из этих черт по отдельности может быть обработана в мозге без задействования осознания.

Даже эмоциональная реакция иной раз происходит подсознательно. Но, когда к делу привлекается сознание, эти компоненты соединяются в единое обогащенное понимание.

Такие наблюдения привели к одной из немногих согласованных точек зрения. Она гласит, что сознание связано с масштабной интеграцией информации в мозге[35]. Как именно взаимодействуют эти два процесса — спорный вопрос. Некоторые теоретики предполагают, что сознание приводит к связыванию информации воедино. Другие выдвигают идею обратной причинно-следственной связи: интеграция информации в сложные сети приводит к возникновению сознания. Самая известная версия — удачно названная теория интегрированной информации Джулио Тонони[36]. Сторонники этой теории вычисляют показатель Ф (фи), который отражает объем интегрированной информации, содержащейся в предмете — будь то мозг, сотовый телефон или что угодно еще. С ростом фи увеличивается и сознание.

Некоторые детали подобных теорий могут не очень хорошо сочетаться с теорией схемы внимания, но их все объединяет общая мысль: сознание связано с интеграцией информации. Здесь я хочу сосредоточиться на существенном касающемся интеграции моменте, который часто упускают из виду. В каком-то смысле сама информация может быть “липкой”, причем некоторые ее виды более липкие, чем другие.

Представьте себе россыпь точек на листе бумаги, большинство из них — черные, а сколько-то — красные. Красные складываются в букву Х. Форма буквы бросается вам в глаза. Точки объединяются

общей цветовой информацией. Важность этого наблюдения впервые поняли гештальт-психологи в начале XX в.: они изучали скрытые правила, согласно которым зрительные образы могут группироваться в единое целое[37]. Само собой, я не имею в виду, что цветовая информация обладает буквальной липкостью или что мелкие атомы информации спонтанно объединяются в более крупные информационные молекулы. Но в определенной информационной системе мозга, на высоком уровне обработки, цветовая информация может слепить вместе разные фрагменты информации о форме. Цвет — классический соединитель информации.

Цвет ограничен зрительной сферой. Он помогает нам группировать предметы в поле зрения, но, конечно, не может объединять информацию от других чувств. А вот пространственное расположение — соединитель более общего уровня. Если вы видите трепетание птицы на ветке и слышите щебетание из того же места, вы, скорее всего, свяжете эти два явления воедино и будете трактовать их как единый объект[38]. Привязывая зрительную и звуковую информацию к одному и тому же фрагменту информации о местоположении, будто соединяя детали конструктора, мозг может объединить весь набор данных в единый, более крупный комплект.

Местоположение работает как гибкий соединитель, в частности, потому, что обладает свойством относительности. Информация о местоположении не описывает конкретных свойств птицы — цвета, формы, громкости чириканья: это бы ограничило универсальность. Она описывает птицу относительно вас: та шестью метрами левее, и это соотношение может одинаково применяться ко всем

конкретным чертам птицы. Ее цвет, форма и издаваемый ею звук — все присутствуют в *одном и том же* месте. Информация о местоположении так хорошо подходит для интеграции других фрагментов данных потому, что в мозге она имеет особый статус координатора. Если бы пространственные сведения каким-то образом оттуда удалили, мир нашего восприятия распался бы на сумятицу разрозненных свойств объектов и чувств[39]. Информация о местоположении — словно волокнистый связующий материал, подмешанный в систему восприятия мозга, который удерживает монолитно все остальное месиво данных. Можно было бы вполне обоснованно сформулировать теорию “интегрированной информации” пространственного расположения — не потому, что интегрированная информация порождает расположение (это ерунда), а потому, что в контексте обработки мозгом сведений о мире информация о местоположении обладает исключительной липкостью.

Хотя данные о местонахождении могут связывать информацию между зрением, слухом, осязанием и, возможно, даже обонянием (например, если вы собака и умеете по запаху определять направление), они все же ограничены в связывающих возможностях. В некоторых сферах информации отсутствует очевидный пространственный компонент: эмоции, мысли, убеждения, математические озарения — ничто из этого не крепится к конкретной точке в пространстве вокруг вашего тела. Данные о местоположении нельзя использовать в качестве *универсального* соединителя.

Реально ли найти в мозге информацию настолько липкую, что она будет работать универсальным

соединителем, связывая любые типы сведений с какими угодно другими?

Роль универсального соединителя в вас может играть схема внимания — информация, описывающая состояние этого внимания. Она имеет отношение к птице, на которую вы смотрите, к звуку, который вы слышите, к мысли, которую вы обдумываете, к эмоции, которую переживаете. Схема внимания описывает взаимные отношения между вами и объектом. В этом смысле внимание напоминает пространственное расположение, которое также является соотношением между вами и предметом. Но, в отличие от местоположения, взаимоотношения с вниманием применимы универсально. Объектом внимания может быть все, что пожелаете, — абстрактное или конкретное, воспринятое или помысленное.

Представьте себе, что птица все еще сидит на ветке слева от вас. Давайте проанализируем конкретный промежуток времени, в течение которого вы уделяете птице внимание. Ваш мозг выстраивает информацию о ее внешнем виде и голосе, о вашей эмоциональной реакции на нее и, может быть, даже глубокую философскую мысль касательно птиц вообще. Вы внимательны к компонентам А, Б, В и т.д. Или же вы в основном вслушиваетесь в птичьи трели и лишь мимолетно рассматриваете перышки птицы, оставляя также капельку внимания на никак с птицей не связанный зуд в руке. В следующий миг ваше внимание может переметнуться куда-то еще, но прямо сейчас вы применяете внимание определенным образом к определенным компонентам своего мира. Чтобы отразить состояния вашего мира, вас самих и соотношения между ними, мозг должен построить внутреннюю модель А, Б, В *и* вашего внимания, а

также связать всё вместе. “Трактуя” внимание как относительное свойство мира, которое нужно моделировать, мозг строит центральный соединитель — схему внимания, к которой неизбежно прикрепляются все остальные комплексы информации в зоне вашего внимания.

Получая доступ к этой переплетенной сети данных, ваше познание узнаёт, что “там птица, она ярко окрашена и красиво поет, я в восторге, нужно потом поискать, как она называется, у меня чешется рука, и все эти компоненты не сами по себе, а объединены по общему знаменателю, потому что я переживаю субъективный опыт — *сознание*, которое овладевает этим всем одновременно”.

Гештальт-психологи в основном имеют дело с чувственным восприятием — изучают цвет, форму, расположение, звук и другие сенсорные характеристики, чтобы понять, как связывается воедино воспринимаемый мир. Теория схемы внимания расширяет гештальт-теорию, добавляя наивысший соединитель. Сознание стягивает все характеристики в единое интегрированное целое — меня, встроенного в мир в данный момент времени. Если данные о местоположении — связующее вещество, оно, наверное, напоминает столярный клей, которым что-то можно склеить, а что-то нельзя. Сознание же больше похоже на универсальный суперклей, связывающий информацию из любых сфер и работающий независимо от того, что попадает в поле нашего внимания. Без него наш цельный мир распался бы на отдельные хаотически плавающие компоненты.

Таким образом, теория схемы внимания предлагает свою версию теории интегрированной

информации. В этой версии сведение информации воедино само по себе не приведет напрямую к возникновению сознания. Но данные в мозге могут быть липкими, а некоторые их виды обладают особой липкостью. Цвет помогает интегрировать информацию в поле зрения. Пространственное расположение может интегрировать данные, поступающие от органов чувств. Скорее всего, и многие другие виды информации обладают липкостью, но различаются ее уровнем. Самая липкая информация в мозге, которая может соединяться с чем угодно, — это сведения об отношениях внимания с объектами в вашем мире. По определению все объекты в зоне вашего внимания разделяют это свойство. В силу такого “сверхгештальт”-качества сознание главенствует над масштабной интеграцией информации в мозге.

Глава

8

Машины, обладающие сознанием

Великий физик и математик сэр Исаак Ньютон пробовал себя и в занятиях алхимией[1]. Судя по тому, сколько он извел бумаги для заметок, он вложил больше сил в провальные попытки получить золото, нежели в успешную попытку открыть закон всемирного тяготения. В одном из рецептов Ньютона значатся Огненный Дракон, Голуби Дианы и Орлы Меркурия: это экзотические названия реально существующих ингредиентов. Так и вижу, как сэр Исаак вместе с другими алхимиками склоняется над пергаментами с неразборчиво записанными рецептами, смешивает и нагревает компоненты, напряженно ждет результата — только чтобы снова разочароваться. Теорий им хватало с запасом, но алхимические теории основывались скорее на аналогиях и мифах, а не на логических построениях. Практикующие адепты не могли сказать: “С учетом того, что нам известно о свойствах материалов, при смешении А и Б должно получиться золото”. Лучшее, что было в распоряжении честолюбивого алхимика, — метод проб и ошибок.

Но я не хочу издеваться над алхимиками. Их ремесленнические пробы и ошибки во многом

положили начало современной химии. А современная наука в результате все-таки раскрыла великий алхимический секрет.

Вот вам рецепт, может быть, захотите попробовать сами. Возьмите элемент тяжелее золота, раскрошите его в ускорителе частиц — и получите несколько атомов драгоценного желтого металла. Или облучите более легкий элемент нуклонами — и постепенно частицы слипнутся и образуют ядро золота. Не знаю, пробовал ли кто-нибудь так делать, но подход этот физически обоснован, а значит, он сработает. Ученые получали элементы и намного тяжелее золота, и более редкие. Выгода от моего способа была бы невелика: на получение каждого атома этого благородного металла уходили бы миллионы долларов, но главное здесь — работающий принцип.

Конструирование машины, обладающей сознанием, как раз и напоминает мне алхимию. Множество научных трудов о сознании полны туманных аналогий и даже мифов. Если насыпать достаточное количество сложности, или обратной связи, или масштабных связей, или таламокортикального резонанса и перемешать все это на огне бунзеновской горелки — проснется ли в машине сознание? Изречет ли она: “У меня есть сознательный опыт и красного, и холодного, и себя самой”?

Теория схемы внимания представляется мне инженерным ответом на алхимическую загадку сознания. Если, руководствуясь этой теорией, построить машину, вложить в нее нужные внутренние модели и дать ей когнитивный и лингвистический доступ к ним, то машина будет обладать потенциальными возможностями, которые вы для нее “наинженерили”. Вам не придется надеяться, что

сознание возникнет от алхимического сплавления компонентов. Машина будет считать, что у нее есть сознание, утверждать, что у нее есть сознание, и говорить о своем сознании — потому что вы встроили в нее конструкт сознания.

Гонка создания машин, обладающих сознанием, начиналась на низких скоростях. Были какие-то пробные попытки, основанные на различных теориях, но не возникло ничего, действительно напоминающего сознание[2]. Энтузиазм мог угасать и от того, что теории сознания чаще сосредоточиваются на метафизическом ощущении, а не на прикладной пользе от сознания. Если вы специалист по информатике и вас интересуют практичные, востребованные продукты — зачем тратить время на метафизику? Теория схемы внимания как раз предлагает некоторые прикладные выгоды. Согласно ей, мозг сформировал конструкт сознания потому, что тот давал два существенных преимущества: во-первых, улучшал внутреннюю регуляцию, а во-вторых, служил основой социального познания.

У нас есть конкретная теория *и* конкретные практические выгоды — теперь искусственное сознание готово к взлету. В каком-то виде оно может появиться уже лет через десять. Первые эксперименты по искусственному сознанию, скорее всего, будут крайне ограниченны. Разумные андроиды вроде Дейты из “Звездного пути” или С-3РО из “Звездных войн”, которые обладают не только сознанием, но и безграничными способностями и блестящим умением поддержать разговор, — дело далекого будущего. Они тоже возможны, но технологии еще не дошли до этого уровня.

Однако, прежде чем пуститься в поэтические рассуждения о грядущем, в котором нас будут сопровождать разумные искусственные компаньоны, мне придется начать с более существенного вопроса. Если мы построим машину, которую станем считать сознающей, — как узнать, что у нее на самом деле есть сознание? Возможно ли удостовериться в его наличии?

В 1950 г. уже упоминавшийся математик Алан Тьюринг предложил способ проверить, способна ли машина мыслить[3]. В придуманном им игровом тесте участвуют три человека: один говорит правду, другой обманывает, третий угадывает, кто из них кто. Они сидят по разным комнатам и общаются только письменно, чтобы не допустить непреднамеренных подсказок.

В начале угадывающий ничего не знает о двух других участниках, кроме того, что один из них — мужчина, а другой — женщина. Задача того, кто говорит правду, — передать угадывающему правильный ответ. Задача обманщика — запутать угадывающего и передать ему неверный ответ. Все трое оказываются вовлечены в сложную информационную войну. Поскольку разрешены любые вопросы и любые ответы, разговор может соскользнуть на социальные двусмысленности и тонкие психологические манипуляции. В конце игры угадывающий должен сделать выбор. Если он угадывает верно, побеждает тот, кто говорил правду. Если он ошибается, побеждает тот, кто обманывал. Я никогда не играл в эту игру и не знаю никого, кто играл бы, поэтому трудно назвать средние показатели успеха обманщика или правдивого участника. Предположу, что у правдивого есть преимущество:

когда вы говорите правду, вас невозможно поймать на противоречии. В любом случае, если бы в эту игру сыграли много раз, можно было бы подсчитать приблизительный уровень успеха разных игроков.

Следующий шаг — заменить обманщика-человека обманщиком-машиной. Если машина сможет выигрывать в эту игру с той же частотой, что и человек, — значит, она может думать, как человек. В этом и заключается изначальный тест Тьюринга.

На первый взгляд его способ кажется слишком усложненным. И уж точно не слишком практичным — чтобы собрать статистику, в игру нужно сыграть много раз. Но, если рассмотреть тест поближе, становится ясно, что это блестящая находка, опередившая свое время. Чтобы хорошо играть, машине нужна модель психического состояния других людей. Более того, ей придется успешно пройти тест ложных убеждений. Машина должна понимать, что другие люди, другие сознания могут содержать в себе ложные убеждения о мире, и ей необходимо постоянно быть в курсе, кто во что верит. Это тест социального познания с очень высоким проходным баллом.

Тест Тьюринга можно считать ранней версией задачи Салли-Энн, которую психологи впервые предложили в 1980-х гг.[4] Я описывал эту задачу в пятой главе, но кратко напомню здесь. В стандартной версии психолог, проверяющий ваше социальное познание, рассказывает вам историю о том, как Энн обманула Салли. Салли положила бутерброд в корзину А и ушла в туалет. Пока ее не было, Энн переложила бутерброд в корзину Б. Когда Салли вернется — в какой корзине она будет искать свой бутерброд? Если вы способны понять, что у Салли есть сознание, что в

этом сознании хранятся убеждения и что ее убеждения могут противоречить реальности, — то вы сможете решить эту задачу. Салли сначала будет искать в корзине A — той, куда она и убрала бутерброд. Какой бы тривиальной ни казалась эта задача взрослым людям с многолетним социальным опытом, дети младше пяти лет — как я уже говорил — с ней не справляются, а кроме людей решить ее могут очень немногие биологические виды[5].

Задача Салли-Энн со всеми ее ментальными моделями и отслеживанием ложных убеждений — на самом деле упрощенная версия теста Тьюринга. Возможно, Тьюринг и думал, что пишет о вычислительных технологиях, но он был также хорошим социальным психологом, на 30 лет опередившим свое время. Чтобы проверить, может ли машина мыслить, я скорее воспользуюсь сложным тестом Тьюринга, нежели поверхностной задачкой Салли-Энн. Тест ставит машину перед более сложной проблемой. Чтобы действовать, как человек, машина должна обладать блестящими языковыми умениями, актуальным знанием повседневной жизни и великолепными способностями логически размышлять об убеждениях других.

Тьюринг никогда не утверждал, что его игра в угадайку — тест на наличие сознания. Он никогда не обсуждал связь этой игры с субъективным опытом, осознанием, квалиа[17], внутренним ощущением, приходящим вместе с обработкой информации, или любыми другими способами разговора о сознании. Если машина изначальный тест Тьюринга проходит, мы можем быть уверены, что это удивительно сложно устроенный компьютер. У него должно быть что-то вроде нормального *содержания* человеческого

сознания, иначе он не смог бы поддерживать разговор, — но не факт, что у него есть собственные сознаваемые переживания.

За годы, прошедшие с исходной публикации Тьюринга, его тест присвоили и переиначили толпы фанатов искусственного интеллекта. Знаменитый тест Тьюринга, о котором все слышали в наше время, довольно заметно отличается от оригинала. И от задачи Салли-Энн тоже отличается. Современная версия сосредоточена не на социальном познании, а на сознании. Чтобы понять, есть ли у машины сознание, достаточно с ней поговорить. Если вы не можете определить, с машиной разговариваете или с человеком, — машина прошла испытание. Новый тест усовершенствован по сравнению с предшествующим — теперь его легко провести с настоящими компьютерами, и это проделывали множество раз. Люди организуют целые конференции, чтобы прогонять машины через осовремененный тест Тьюринга; самая известная из них — это ежегодное соревнование компьютерных программ за премию Лёбнера (и все же до сих пор ни одна машина не смогла повсеместно убедить ученых, что у нее есть сознание).

Но даже если бы какая-то машина прошла этот современный тест Тьюринга, мы по-прежнему не смогли бы определить, есть ли у нее внутренний опыт. Тест способностей к разговору нелогично применять в качестве теста на наличие сознания. Учитывая, сколь легко убедить некоторых в том, что сознание есть у рек и деревьев, может оказаться совсем нетрудным одурачить хотя бы несколько человек и заставить их считать обладающей сознанием машину, у которой на самом деле сознания нет. Если бы мы построили машину, которая действительно обладает сознанием,

даже она могла бы провалить тест Тьюринга. Сознание не всегда означает способность вести светскую беседу. О чем говорить — да трехлетний ребенок не прошел бы этот тест! Как и некоторые взрослые. Как и домашние собаки, хотя львиная доля человечества вполне уверена, что у собак есть осознаваемый опыт. Просто тест не очень работает. Но сама идея теста Тьюринга настолько вросла в нашу культуру, мы так с ней сроднились, что будто бы приняли ее в качестве работающего решения проблемы. Обоснование этого обычно звучит примерно так: “Я знаю, что у меня есть сознание, поскольку имею непосредственный опыт переживания своего собственного разума. Но я не могу знать, сознательны ли другие люди. Как бы мне ни хотелось в это верить и как бы я ни любил своих детей, жену и кошку, я никогда не смогу напрямую испытать опыт их сознания. Это невозможно. Придется удовлетвориться лучшим из того, что мне доступно, — я условно *допущу* наличие у них сознания, поскольку они ведут себя так, будто оно у них есть. Аналогичным образом я не могу доказать наличие сознания у машины. Это недоказуемо. Но с тестом Тьюринга я могу сделать лучшее из возможного — проверить, принадлежит ли машина к той же категории, что мои жена, дети и домашние животные. Я смогу определить, способна ли машина вести себя настолько убедительно, что уподобится сознающему существу, — тогда мне придется условно допустить, что она действительно обладает сознанием”.

Так выглядит типичное рассуждение. Но я бы хотел развеять таинственный ореол вокруг теста Тьюринга. Исходя из теории схемы внимания, мы *можем* выяснить с объективной достоверностью, обладает ли

машина тем же типом сознания, что и люди. Причем прямой личный опыт — *не* единственный способ (это вообще не слишком хороший способ) выяснить что-то о собственном сознании.

“*Конечно*, у меня есть сознание! Я точно знаю это — у меня есть непосредственный опыт его переживания”.

Если это не круговая логическая ловушка, то что же? Сознание *является* непосредственным опытом. Следовательно, данное утверждение эквивалентно такому: “Я знаю, что у меня есть сознание, потому что у меня есть сознание”. Как я уже говорил, машина — заложник содержащейся в ней информации. Определенная внутренняя модель сообщает машине, что у нее есть сознание, и получается, будто она “знает”, что у нее есть сознание. Внутренняя модель сообщает машине, что ее сознание не имеет физического носителя и является сугубо частным опытом, — получается, машина “знает”, что ее сознание не может подтвердить никто другой.

Но внутренняя модель — это информация, а информацию можно объективно измерить. Нам нет нужды полагаться на личные утверждения. В теории схемы внимания, чтобы определить наличие сознания у машины, нам нужно “ощупать ее внутренности” и понять, содержит ли она схему внимания, а затем прочесть информацию в этой схеме внимания. Тогда мы с объективной достоверностью выясним, может ли эта машина, подобно нам, полагать, что у нее есть осознаваемый субъективный опыт. Если в ее внутренней модели есть соответствующая информация, то да. Если нет — то нет. Все это, в принципе, можно измерить и подтвердить.

Измерить имеющуюся в человеческом мозге информацию сложно, но не невозможно. В

ограниченных масштабах ученые уже этим занимаются. Хороший набор электродов на вашей голове поможет исследователям предсказать, куда вы собираетесь двинуть рукой — вправо или влево[6]. По качественной магнитно-резонансной томограмме [18] зрительной коры ученые могут определить, смотрите вы на лицо или на дом[7]. Считывать более сложную информацию пока невозможно, но это решается усовершенствованием технологий. Уверен, что технологии считывания данных из мозга в действительно высоком разрешении разовьются со временем (и это внушает тревогу). Возможность измерять информацию в человеческой схеме внимания должна существовать в принципе, пусть даже извлечение таких сложных сведений из мозга — дело далекого будущего. Главное здесь то, что в теории схемы внимания сознание не обречено навеки оставаться исключительно частным и личным опытом. "Я знаю, что у меня есть сознание, но никогда не смогу узнать, есть ли оно у тебя" — это неправда. Узнать, считает ли чей-то мозг, что у него есть сознание, такое же как у меня, — это вопрос развития технологий для считывания из него информации.

Если вместо биологического мозга перед нами искусственный интеллект, то должно быть намного проще "заглянуть под капот" и измерить сознание. Инженерам нужна будет возможность подключиться к сделанной человеком машине и измерить ее информационное содержание — раз уж у нас есть для этого инструменты много лучшие, чем для подключения к мозгу (да к тому же у кого-нибудь наверняка найдутся чертежи этой машины).

Путаница начинается, когда люди спрашивают: "Я понимаю, что вы можете измерить информацию, но

как вы определите, есть ли у машины то самое внутреннее чувство? Только машина может переживать опыт своего сознания, а раз так, кто, кроме машины, может о нем знать?” Такой вопрос возникает, когда ваше мышление спрашивает совета у вашей же схемы внимания. Она говорит вам, что вы обладаете личным, глубоко индивидуальным, нематериальным ощущением, доступ к которому есть только у вас. Но наша гипотетическая машина содержит все те же элементы, что и вы, — в том числе схему внимания, содержащую подобную информацию. Посоветовавшись со своей схемой внимания, машина тоже “поймет”, что у нее есть глубоко индивидуальное, личное, нематериальное ощущение. И вы, и машина созданы одинаковым образом и попадаетесь в одну и ту же логическую ловушку. Вы знаете лишь то, что знаете, и можете сообщать лишь ту информацию, что содержится в вас, — но эту информацию теоретически может считывать и кто-то другой.

Я хочу сказать, что на вопрос, не имеющий, казалось бы, ответа, — как мы можем понять, есть ли у машины сознание? — в принципе, может найтись достоверный ответ, когда у нас будет правильное оборудование для добычи информации. Нам нет нужды полагаться на тест Тьюринга, который в лучшем случае действует не напрямую, а в худшем — вообще не имеет отношения к проблеме сознания.

В оставшейся части этой главы я рассмотрю практические соображения постройки обладающей сознанием машины с использованием теории схемы внимания в качестве руководства. Нам нужны четыре составляющие. Во-первых, у машины должно быть искусственное внимание — способность сосредоточивать свои ресурсы и переключаться,

глубоко обрабатывая сначала один предмет, затем следующий. Во-вторых, машине требуется схема внимания — внутренняя модель, которая в общих чертах описывает внимание и таким образом сообщает машине о наличии субъективного сознания. В-третьих, машине необходим адекватный диапазон информационного содержания. У нас может быть машина, которая осознает, например, только зрительную информацию (скорее всего, первая попытка будет именно такой), но со столь ограниченным диапазоном эта машина вряд ли сможет сойти за человека. В идеале она должна располагать намного более обширным содержанием. В-четвертых, машине нужна сложно устроенная поисковая система, которая умеет обращаться к внутренним моделям и беседовать о них, — чтобы мы могли поговорить с машиной и она была бы в состоянии сообщить нам о своем сознании. Если мы сможем выстроить эти четыре компонента, у нас будет машина, обладающая чем-то вроде человеческого сознания.

Давайте коротко рассмотрим эти компоненты по очереди и выясним, насколько они реализуемы на сегодняшний день.

Первый компонент — внимание. Машина должна уметь сосредоточивать ресурсы обработки информации и переключаться с предмета на предмет — с яблока на пончик, с него — на человека в углу комнаты, а с того — на внутреннее явление, например всплывшее воспоминание.

В каком-то виде искусственные устройства уже обладают вниманием. Планомерное управление ресурсами — стандартное свойство любого современного компьютера. К тому же многие

исследователи описали или построили вычислительные модели — компьютерные симуляции, — которые воспроизводят различные аспекты внимания, подобного человеческому[8].

Но чтобы сделать сознающую машину, нам потребуется особый вид внимания — однако этого еще никто не достиг. Человеческое внимание обладает специфическими свойствами. Будучи сосредоточенным на предмете, оно позволяет подробно обработать этот предмет и извлечь из него глубинное значение. Использование внимания помогает вам определить аффордансы предмета: можно ли его схватить, пнуть, укусить? Внимание предоставляет и возможность действовать, а именно — принимать решения о том, что делать с предметом. Оно позволяет сохранить его в памяти и вернуться к нему позже. Динамика внимания означает, что более громкие и яркие стимулы умеют перехватывать фокус обработки информации, но в то же время внимание не полностью подчинено капризам среды — оно подчиняется и вам, причем так, что внутреннее указание может перенести внимание с предмета на предмет. А главное, внимание способно легко переключаться между разными модальностями информации. К примеру, оно может направляться на определенные места в пространстве (это основное свойство внимания, воплощенное в искусственных системах на данный момент), либо же сосредоточиваться на цвете, движении, вкусе, осязании давления или даже на чем-то внутреннем — воспоминании или мысли.

В мире искусственного интеллекта я до сих пор не встречал ничего подобного такой сложности внимания. Одна из причин в том, что искусственный

интеллект, как правило, сосредоточен на одной модальности. Возьмем, например, систему искусственного распознавания лиц: это важная функция, которую можно применять различными способами, но в основном она обходит вопрос внимания стороной. Направить камеру на лицо, которое вы хотите идентифицировать, умеет и простая версия пространственного внимания, а затем свою работу делает уже алгоритм распознавания. Данной специфической машине в принципе не требуется переключать фокус внимания между различными модальностями информации — с лиц на звуки, со звуков на мысли: она работает исключительно в границах сферы распознавания лиц. В сущности, эта система не имеет отношения к вниманию, разве что в самом тривиальном, очевидном смысле наведения взгляда на цель. Искусственное внимание технически возможно, и многие специалисты работают над ним, но до сих пор оно не обрело межмодальной связности, подобной той, что присуща человеческому вниманию. Первый компонент сознающей машины не стоит на месте, но ему предстоит еще развиваться и развиваться.

Второй компонент, необходимый для машины, — схема внимания, ключевая внутренняя модель, которая в общих чертах описывает внимание и таким образом сообщает машине о наличии сознания.

Один мой коллега как-то сказал, что у любого стационарного компьютера давно уже имеется что-то вроде схемы внимания — комплекс информации, который отслеживает распределение ресурсов обработки. Получается, что, согласно моей теории, любой компьютер уже обладает сознанием. Другой коллега пообещал запросто запрограммировать всю мою теорию за полдня. Он мог бы состряпать

искусственную версию зрительного внимания, подсоединить ее к модулю, который отслеживает его изменения, — и вуаля, если я прав, устройство должно начать осознавать себя. Группа коллег и в самом деле создала устройство искусственного внимания со схемой внимания, которая помогала этому устройству отслеживать внимание и управлять им[9]. Как и ожидалось, внутренняя модель поспособствовала тому, чтобы сделать управление вниманием более эффективным, — но стала ли машина осознавать себя?

Я практически уверен, что нет. Ни один из приведенных примеров не свидетельствует о наличии сознания. Схема внимания — не магический талисман, который можно засунуть в компьютер, чтобы оно возникло. Это пакет информации. Если пакет несет простое сообщение вроде: “Переменная А — в допустимых пределах” и если машина может вербализовать эту информацию — она не скажет вам: “Я переживаю сознательный опыт!”, она расскажет только о переменной А. Во взаимоотношениях схемы внимания с сознанием нет никакой магии. Если схема внимания сообщит машине о наличии сознания, то машина получит информацию о том, что у нее есть сознание.

Человеческая схема внимания формировалась сотни миллионов лет в ходе эволюции, и ее содержание — довольно странное, биологически неоднородное, “комковатое”. Она описывает внимание как незримое свойство, как разум, который может переживать или вбирать в себя объекты, как силу, которая позволяет человеку действовать и помнить, как нечто не имеющее собственного физического воплощения, но существующее внутри

человека в виде его личного опыта. Схема внимания — не просто указатель на предмет, не просто пара строк кода. Она рисует глубокую картину внимания и его предсказуемых последствий. Постройте машину, обладающую схемой внимания *такого же типа*, содержащую странные, биологически беспорядочные “комки”, как у настоящей, — и получите машину, которая подобным человеку образом сможет утверждать, что обладает сознанием. Для построения искусственной схемы внимания нет принципиальных технологических препятствий. Вероятно, это самый простой (и уж точно самый подробно описанный) из всех необходимых в машине компонентов.

В качестве следующего компонента сознающей машине понадобится содержание. Пожалуй, конструирование богатого и разнообразного содержания сознания — самая сложная задача, поскольку она не имеет конечного решения. Ирония состоит в том, что описанная выше пресловутая трудная проблема (сделать так, чтобы машина хоть что-то осознавала), скорее всего, окажется самой простой, а простая проблема (дать машине широкий диапазон материала для осознания) может оказаться самой трудной. Думаю, что усилия, направленные на создание искусственного сознания, начнутся с обработки сенсорной информации, особенно зрительной, — поскольку мы очень много знаем о работе сенсорных систем в мозге и об их взаимодействии с вниманием. Но мне кажется, что все мы отнесемся с некоторым скепсисом к машине, которая сможет осознавать черную точку на экране и ничего больше. Это будет полезным лабораторным тестом, но мы не станем считать такую машину функционально сознающей.

Даже если нам удастся заставить машину осознавать широкий диапазон сенсорной информации, в конечном итоге она должна суметь выйти за границы чувств и включить в свое сознание абстрактное мышление. И здесь инженерные проблемы усугубляются. Мы не очень много знаем о том, как абстрактное мышление взаимодействует с механизмами внимания и осознания. Мы знаем, что человек может обращать внимание на определенную мысль, удерживать ее в уме и сосредоточивать на ней ресурсы за счет других возможных мыслей — и даже оттягивать на нее внимание от обработки сенсорной информации[10]. Нам всем случалось так задумываться, что мы переставали замечать происходящее вокруг нас, во внешнем мире. Но наука о мозге еще не разобралась, как внимание к абстрактным мыслям работает на механическом уровне. Может быть, только через десятки лет мы сможем придумать, как построить машину, обладающую такими способностями.

Но подозреваю, что самым крепким орешком окажутся эмоции. Эту область информации в мозге мы понимаем хуже всего. Самое большее, что можно о ней сказать, — это что мы определили местоположение некоторых связанных с эмоциями мозговых структур[11]. Иногда люди называют их эмоциональной сетью, или эмоциональной кругом, но детальное понимание функционирования этой системы отсутствует.

Одна из важнейших в этом смысле областей мозга — гипоталамус, структура размером с лесной орех в основании мозга. Более ста лет назад швейцарский психолог Вальтер Гесс открыл, что, если ввести в гипоталамус электрод и для активации нейронов

пустить по нему очень слабый ток, можно вызвать то, что для стороннего наблюдателя выглядит как определенные эмоциональные состояния[12]. Его эксперименты воспроизводились множество раз[13]. Страх, гнев, похоть, голод — всех этих состояний удается добиться, стимулируя различные точки в гипоталамусе.

Еще один центральный процессор эмоций в мозге — миндалина, структура размером с ядро этого ореха. Их две, по одной в каждом полушарии мозга, и они активируются, когда мы переживаем эмоции, особенно когда определенная эмоция ассоциируется с конкретной ситуацией или зрительным образом[14]. Когда вы злитесь, глядя на ненавистного политика, это именно миндалина связывает зрительное восприятие с эмоциональным переживанием. Еще одна область мозга, играющая важную роль в эмоциональном опыте, — орбитофронтальная кора, это самая нижняя часть префронтальной коры, расположенная за костью глазниц. Она участвует в принятии решений, основанных на эмоциональном содержании[15]. Все эти факты нейронауки удивительны, но очень мало говорят нам о том, как построить искусственные эмоции или как их встроить в механизмы внимания и осознания.

Подсказку можно найти в опыте, который есть у любого человека, — переживании эмоций вне сознания[16]. Мы обычно не рассматриваем их как нечто, существующее отдельно от сознания, но такое все же возможно. Вот пример из повседневной жизни — вы ощущаете, скажем, напряжение, возбуждение, стресс или гнев, но не слишком сильные, и не осознаете их. И только когда друг скажет вам: “Ну и состояньице у тебя сегодня!”, вы перенаправите свое

внимание — и осознание — и поймете: “Ничего себе, а я и *правда* напряжен”. Кто-то осознает свои эмоции лучше, кто-то — хуже, но всем случается генерировать их, не осознавая того. Любая теория эмоционального опыта должна учитывать такое расхождение между эмоциональным состоянием и осознаваемым переживанием опыта.

Нейробиолог Джозеф Леду, основываясь на своих новаторских изысканиях, касающихся обработки эмоций в мозге, сформулировал элегантное представление об их осознании[17]. Эмоциональные состояния организованы вне сознания, в структурах в глубине мозга, ниже коры — в только что упомянутых гипоталамусе и миндалине. Для осознания своего эмоционального состояния нам необходимо, чтобы данные из этих глубинных структур достигли корковых систем, в которых они интегрируются с когнитивной информацией о сознании. Согласно этой гипотезе, осознание эмоций состоит из двух частей: данных, определяющих эмоциональное состояние, и данных, определяющих сознание. Мозг может обработать сложный комплекс информации “я осознаю свою эмоцию” точно так же, как “я осознаю яблоко”.

Но, даже пользуясь этим представлением, проектирующий нашу машину инженер не сможет ответить на вопрос: “Какая информация определяет эмоции?” Ответа и впрямь не знает никто.

Но и здесь есть подсказка. Ее можно обнаружить в знаменитой теории эмоций, которую предложили психологи XIX в. Уильям Джеймс и Карл Ланге[18]. Согласно теории Джеймса — Ланге, эмоция начинается с телесного ощущения. Сердцебиение ускоряется, желудок выделяет кислоту, кожа

покрывается холодным потом — и тогда, заметив эти телесные изменения и изучив контекст, мозг создает сюжет: “Я встревожен” или “Я взволнован”. Одна из лучших демонстраций этого эффекта — известный эксперимент, проведенный в 1970-х гг. на мостах в канадском Ванкувере[19]. В данном исследовании женщина обращалась к пешеходам-мужчинам с просьбой ответить на несколько вопросов. Некоторых мужчин останавливали прямо в середине шаткого подвесного моста над глубоким каньоном. Других — на небольшом и устойчивом деревянном мосту. Потом испытуемых спрашивали, насколько привлекательной им показалась женщина-интервьюер. Мужчины, которых останавливали на опасном мосту, оценивали женщину как более привлекательную, а те, что отвечали, стоя на надежной конструкции, — как менее. Предположительно, подвесной мост вызывал ускорение сердцебиения и холодный пот — и эти изменения ошибочно приписывались сексуальному возбуждению.

Современная психология отчасти признает теорию Джеймса — Ланге. Некоторые составляющие эмоций коренятся в физических ощущениях тела, но другие связаны со сложными высокоуровневыми репрезентациями, и эти компоненты взаимодействуют замысловатым и плохо изученным образом. Хотя упомянутая теория не полностью соответствует действительности, она поднимает ряд любопытных вопросов об эмоциях машин. Понадобятся ли андроиду желудок, потовые железы и сердце, чтобы переживать эмоции подобно человеку? Могут ли возникать настоящие эмоции у машины, если она никогда не узнает вкус пищи и не прочувствует пищеварение, а в ее теле ни разу не поднимется адреналиновая волна, заставляющая

бороться или убегать? Лично я подозреваю, что в машине могут происходить процессы очень похожего характера, пусть и не полностью идентичные человеческим эмоциям. Возможно, чтобы дать эмоциям машины физический субстрат, придется оснастить ее сенсорами по всему телу.

В первых опытах искусственного сознания, скорее всего, убедительных эмоций ждать не придется. Может быть, машинам удастся воспроизвести эмоциональную окраску голоса, но им потребуется еще немало времени, чтобы освоить более глубокий уровень эмоций и способы их отображения в человеческом мозге. А пока что нас ждут гораздо более простые задачи. Голливудский стереотип безэмоционального андроида еще на какое-то время останется верным.

Четвертый и последний компонент, без которого не обойтись сознающей машине, — говорящая поисковая система. Мы хотим создать такое устройство, которое сможет болтать с нами о своем сознательном опыте. Строго говоря, способность разговаривать не является необходимой для сознания, но, мне кажется, большинство людей полагают, что цель искусственного сознания — машина, подобная человеку по своей способности говорить и понимать. Мы хотим, чтобы она могла поддержать беседу.

Представляется, что эта часть проблемы уже решена — у нас есть цифровые помощники, например Сири и Алекса. Они могут понять произнесенный вслух вопрос, порыться в базе данных и ответить. Разве этого недостаточно, когда мы хотим, чтобы машина вела разговоры о своих внутренних состояниях? Но проблема наша обманчива и каверзна. Сири в основном действует в языковой сфере. Вы

скармливаете ей слова, она ищет в интернете другие слова и в ответ выдает опять же слова. Но вы не получите ответа на вопрос, где ближайший ресторан — Сири не знает, что такое ресторан, кроме как в составе статистической группировки слов. А человеческий мозг может переводить речь в невербальную информацию и обратно. Если кто-то спросит вас: “Каков вкус лимона по сравнению с апельсином?” — вы ответите не так, как интернет-поисковик. Ваш ответ будет основан не на словесных ассоциациях. Вы переведете речь во вкусовую информацию, вспомните два вкуса, сравните их и переведете ответ обратно в слова. Такой банальный для нас перевод туда-обратно между речью и другими сферами информации невероятно трудно воспроизвести искусственно. Насколько мне известно, для этой задачи еще не нашли общего или системного решения. Гугл способен до некоторой степени переводить информацию из зрительных образов в слова, но нашей сознающей машине придется сопоставлять информацию из всех возможных сфер.

Все это звучит весьма многообещающе, но путь очень сложен — так насколько же близко мы подошли к сознающим машинам на самом деле?

Первые образчики машин, обладающих сознательным восприятием зрительной информации, — в которых есть и внимание, и схема внимания, причем оба элемента направлены на визуальную информацию, — могут быть построены уже в ближайшие десять лет. Думаю, что для создания машин, содержащих столь же обширный диапазон разнообразной информации, как в человеческом сознании, потребуется больше времени. Изготовление устройства, которое способно видеть, слышать, осязать, чувствовать вкус, мыслить абстрактно и

производить эмоции, которое умеет координировать эти сферы между собой и внутри каждой из них при помощи единого интегрированного фокуса внимания и которое может говорить обо всем этом диапазоне содержаний, — весьма долговременный проект. Буду удивлен, если мы получим что-то хотя бы приближающееся к этому в течение следующих 30 лет (мой прогноз — это займет порядка 50 лет). Но мне и раньше случалось удивляться, а технологии развиваются с ошеломляющей скоростью.

Люди мечтали о разумных роботах с тех пор, как появились механические технологии. В “Илиаде” Гомера, написанной почти 3000 лет назад, упоминаются самодвижущиеся треножники, которые изобрел бог Гефест. Пятьсот лет назад Леонардо да Винчи разработал и, возможно, построил механических роботов для развлечения своих покровителей[20]. Сегодня всем нам знакомы HAL 9000 Стэнли Кубрика из фильма “2001: Космическая одиссея” и C-3PO Джорджа Лукаса из “Звездных войн”. Относительно недавний фильм Алекса Гарленда “Из машины” целиком посвящен тому, как люди могли бы взаимодействовать с разумными роботами. Обладающие сознанием машины так долго занимали умы людей, что, если увидеть, сколь близко мы подошли к их появлению в нашей жизни, становится жутковато. Вспоминается поговорка: “Будь осторожен в своих желаниях”.

Пока в нашем распоряжении нет готовой культурной модели того, как нас могут изменить сознающие машины. Научная фантастика зачастую предвосхищает технические приспособления будущего, но предсказать их социальные последствия оказывается несколько сложнее. Например, никто не

мог предвидеть, какой переворот в обществе произведут сотовые телефоны. Вначале мы думали, что они просто заметно облегчат нам жизнь — подобно микроволновым печам, но постепенно они превратились в третье полушарие человеческого мозга. Смартфоны перестроили наш политический, экономический и социальный мир. Поэтому вот единственный прогноз, который я могу сделать с абсолютной уверенностью, — никому не удастся столь же уверенно предсказать социальные последствия появления машин, обладающих сознанием.

Наша научная фантастика пытается нарисовать картину будущего с искусственным сознанием, но сомневаюсь, что эта картина окажется сильно приближена к реальной. В “Звездных войнах” сознающие машины, похоже, не добавляют обществу слоя этических проблем. Они всего-навсего второсортные граждане, обслуживающий персонал, покорные и умелые рабы, милые, но легко заменимые — с ними обычно обращаются даже менее уважительно, чем с домашними животными. Следует ли относиться к разумным машинам именно так — или же им полагаются моральные права? Эту проблему и затрагивает знаменитая повесть Айзека Азимова “Двухсотлетний человек”[21]. Раз разумные машины создаем мы, есть ли у нас право управлять ими и убивать их? Тот же этический вопрос поднимается в фильме “Бегущий по лезвию”, снятом по рассказу Филипа Дика[22]. Писатели, режиссеры, ученые и философы ломают головы над этими вопросами, не находя ответов[23]. Чтобы разработать прикладные этические правила, нам, возможно, придется дождаться постройки поистине человекоподобного искусственного сознания. Проблемы, скорее всего, будут выглядеть не так, как

их описывает научная фантастика — в ней обычно замаскированы размышления о трудностях, стоящих перед нами сегодня. Трудности завтрашнего дня могут оказаться неузнаваемо другими.

Пожалуй, главная грозная неопределенность состоит не в том, хорошо или плохо мы будем обращаться с сознающими роботами, — а в том, как будут обращаться с нами они. Обладающие сознанием роботы стали типичными для апокалиптических сценариев. В “Терминаторе” Скайнет пытается всех нас убить. В “Матрице” машины порабощают нас. Но опять же — это все научная фантастика, не реальность.

Может быть, из-за того, что мы переели научно-фантастических историй, мало кто понимает идею сознающей машины правильно. Большинство предполагает, что сознание будет включаться в машине по щелчку тумблера: машина внезапно пробудится, начнет осознавать себя как независимого деятеля, отдельного от остального мира, и немедленно устремится к удовлетворению своих собственных интересов. Она может начать убивать людей, потому что они конкурируют с ней за ресурсы. Машины с такими способностями привели бы к трагедии для человечества. Но плохая новость на самом деле вот в чем. Сверхразумные, автономные, активные машины, которые принимают решения и учитывают свои интересы, уже здесь (по крайней мере, их первые представители), и они стремительно умнеют с каждым днем. Им не нужно сознание, чтобы прийти к выводу, что люди — препятствия на их пути, и поубивать нас. Беспилотный грузовик, имеющий собственную программу действий, может переехать вас — и вы будете в любом случае мертвы независимо от того,

получит ли грузовик субъективный опыт хлюпающего звука под его шинами.

Для меня как исследователя сознания вопрос стоит так: что произойдет, если мы добавим дополнительный элемент — модель себя, которая сообщит машине, что подобно тому, как у нее есть субъективный опыт, у других он тоже может быть? Как это изменит социальный баланс?[24] Думаю, что после этого будущее с разумными машинами станет более обнадеживающим.

Люди — всецело социальные животные. Это определяет нас как биологический вид. Но у человека не может быть социальной компетентности без понимания того, что такое сознание, и без способности интуитивно приписывать его другим людям. Способность эта позволяет нам признавать разумность других, понимать мысли и чувства друг друга, тщательно подбирать верные реакции. Это объединяющий людей социальный клей, корневая система социума и сотрудничества.

Представьте себе мир, в котором люди утратили это глубинное, основанное на моделях знание о сознании. Вместо него в нас встроили альтернативную схему. Мы не впали в безразличие подобно людям с тяжелыми повреждениями мозга. Мы все так же в состоянии действовать как автономные существа, учиться, учитывать свои интересы и принимать разумные решения, но все это мы делаем без малейшего понятия о сознании. Мы становимся независимыми сущностями, у каждой из которых есть собственные интересы, а с другими людьми обращаемся как с объектами. Нам не удается по-настоящему сотрудничать, хотя и хватает ума догадаться, что сотрудничество дает преимущества, —

но мы больше не способны понимать, что у других людей есть сознание, а значит, и не можем угадывать психические состояния друг друга или координировать наши мысли, цели и действия. Убийство становится столь же обыденным, как отодвигание с дороги препятствия, потому что теперь мы неспособны понять ценность чужого сознания. Мы превращаемся в цивилизацию умных, целеустремленных монстров.

Так выглядит мир несознающих, но умных машин, и этот мир мы создаем прямо сейчас. Я бы предпочел жить в мире машин, которые знают, что такое сознание, и могут приписать его и мне, и другим людям, — равно как сегодня я благодарен за то, что живу в мире, где сознание могут приписывать друг другу люди.

Я не утверждаю, что сознание — чудотворный компонент, и стоит только добавить его в машину, как она станет вести себя этично. Всем известно, что среди нас есть и люди-человеконенавистники, социопаты и агрессоры. Все они обладают сознанием и, наверное, знают, что те, кому они причиняют боль, тоже сознательны. Но это потерянные кусочки огромной мозаики. Они всегда в меньшинстве, поскольку человечество по сути своей — просоциальный биологический вид. Каждый раз, входя в супермаркет, я полагаюсь на то, что окружающие будут более-менее соответствовать социальной матрице. Эволюция дала нам несовершенное, но статистически довольно неплохое решение проблемы социальной слаженности. Именно наше гиперсоциальное мышление удерживает человечество на плаву и отличает нас от менее сговорчивых и более враждебных биологических видов. У этой способности, скорее всего, много разных источников,

но она была бы невозможна без внутренней модели, сообщающей нам, что сознание есть и у нас, и у других.

Раз уж мир гонится за созданием машин, подобных человеку разумом и автономностью, а то и превосходящих его по этим параметрам, мне кажется, нам стоило бы дать этим машинам и какие-нибудь очеловечивающие их черты, чтобы у них был шанс интегрироваться с нами. Почему бы не попробовать то же решение проблемы социального сотрудничества, которое эволюция нашла для нас? Если теория схемы внимания верна, то создание искусственного сознания — один из самых полезных шагов, которые мы можем сделать, чтобы снизить технологические риски, грозящие нам в будущем.

Глава

9

Перенос личности на искусственные носители

Вечно жить я не желаю. Мне всегда казалось, что если в жизни есть смысл, то он должен заключаться в том, чтобы внести свой полезный вклад в общество, а если жить вечно, то мы скорее станем забирать, нежели вкладывать. А еще станет скучно. В компьютерной игре, если не рискуешь погибнуть, пропадает все ощущение опасности, весь азарт — и в душу вползает апатия. Ну, мне так кажется. Может быть, на склоне лет я пересмотрю свое сегодняшнее мнение. Но правда в том, что многие люди — может быть, большинство — хотят жить неопределенно долго, и технологии развиваются в этом направлении. Я не имею в виду медицинское бессмертие, которое, судя по всему, физически недостижимо. Я имею в виду оцифровку психики — перенос личности, сущности человека, содержания мозга на искусственный носитель. Рано или поздно люди изобретут посмертие, и тогда человеческий разум сможет жить неопределенно долго в имитации Вселенной и станет взаимодействовать с реальным миром при помощи видеосвязи или робототехнических приспособлений. Жизнь полна странностей, но из всех странных вещей, с которыми я столкнулся в окружающем меня мире

или в своих мыслях, страннее всего, пожалуй, перспектива компьютерного посмертия.

Если верна теория, над которой я работал последнее десятилетие, то все содержание психического мира — воспоминания, эмоции, личность, даже само сознание — результат работы физических механизмов мозга, который можно скопировать. Нам нужно научиться сканировать мозг в достаточном разрешении, чтобы создать его имитацию, искусственный дубликат его данных и алгоритмов, который сможет пережить биологическую смерть человека. Я отнюдь не хочу обесценить этим утверждением сложность человеческого устройства. Как я уже говорил, разум — это скульптура из триллионов информационных нитей, постоянно меняющаяся и прекрасная в своей сложности. Но в ней нет ничего настолько загадочного, чего нельзя было бы скопировать на другой информационный носитель — как сейчас это можно сделать с файлами для переноса их с одного компьютера на другой.

Многие специалисты считают, что технология переноса личности на искусственный носитель уже на подходе[1]. Мне сроки видятся не столь радужно. С одной стороны, все связанное с информационными технологиями развивается с поразительной скоростью. С другой стороны, все связанное с пониманием и сканированием устройства мозга продвигается намного медленнее. И все же неизбежно, что в один прекрасный день мы изобретем эту технологию. Нам могут потребоваться сотни лет или, напротив, это случится раньше, чем я ожидаю, но в целом технологические тенденции и человеческие устремления указывают в этом направлении.

Описанная задача принципиально отличается от создания искусственного сознания. Она требует меньших знаний о том, как взаимодействуют друг с другом различные компоненты мозга, поскольку не нужно конструировать их с нуля. Все, что от вас требуется, — скопировать уже существующий мозг. Нет особой нужды разбираться, почему он устроен так, как устроен, — достаточно сделать точную его копию. Главная сложность таится в том, что копирование мозга требует невероятно высокого уровня детализации при сканировании.

Чтобы построить успешно функционирующую машину по переносу психического мира, сначала нужно определить в мозге минимальный набор данных, которые содержат суть личности. Большинство нейробиологов считает, что обработка информации в мозге выполняется в основном нейронами, которые соединяются друг с другом синапсами — специализированными контактными участками, позволяющими информации идти управляемым потоком от нейрона к нейрону. Человеческий мозг содержит порядка 86 млрд нейронов[2]. В нем может находиться примерно 100 трлн синапсов, а возможно, даже на порядок больше.

Представление о том, что мозг работает посредством нейронов и синапсов — так называемую нейронную доктрину, — сформулировал чуть больше века назад испанский ученый Сантьяго Рамон-и-Кахаль[3]. Он был одним из гениев нейробиологии и в 1906 г. получил Нобелевскую премию по физиологии[19]. Исследователь окрашивал мозговую ткань и изучал ее срезы под микроскопом, таким образом он смог отследить сложные переплетения нейронов и получить первое истинное представление

о том, как работает мозг[4]. Информация течет от нейрона к нейрону через их тончайшие дендриты и терминали. Движение информации управляется синапсами между нейронами. Где-то они останавливают поток, где-то пропускают, направляют по определенным путям и сетям в мозге — это происходит на всех этапах прохождения информации: от входных данных через внутреннюю обработку к результату на выходе. В общих чертах представление испанского ученого не отличается от того, которым пользуются современные нейробиологи[20]. Изящные рисунки Кахаля, на которых изображены отдельные клетки мозга, до сих пор приводятся в учебниках.

Вдохновившись нейронной доктриной, инженеры и ученые создали искусственные нейроны и соединили их в обширные сети, чтобы посмотреть, насколько обучаемой и умной окажется имитация нервной системы[5]. Эта технология перевернула наш мир. Искусственные нейронные сети оказались невероятно адаптивными и обрели огромную мощь. Поисковые системы в интернете, цифровые помощники, которые будто бы понимают речь, беспилотные автомобили, алгоритмы торгов на бирже Уолл-стрит, начинка вашего смартфона — все эти ставшие обыденными элементы нашего времени работают в том числе на искусственных нейронных сетях.

Сантьяго Рамон-и-Кахаль — не только отец современной нейробиологии, но и (чего никто не мог предположить) основоположник сегодняшней технологической революции.

Лежащий в основе нейронных сетей принцип состоит в том, что каждый отдельный нейрон предельно прост, но, когда громадное их количество

соединяется вместе, они становятся огромной вычислительной мощью. По сути нейрон всего лишь посылает сигнал. Предположим, нейрон А соединен синапсом с нейроном Б. Когда А активируется, он посылает электрический сигнал, который идет по всей его длине, доходит до синапса, перескакивает дальше с помощью химического переносчика и воздействует на Б. Если это возбуждающий синапс, то сигнал усилит имеющуюся активность нейрона Б, для которого тогда повысится вероятность послать свой собственный сигнал. Если это тормозной синапс, то сигнал, перескочивший по нему, утихомирит активность нейрона Б, для которого тогда понизится вероятность послать свой собственный сигнал. Также синапсы могут различаться по силе воздействия — некоторые передают более мощные сигналы, позволяя нейрону А сильнее воздействовать на нейрон Б, а некоторые слабее, и тогда нейрон А почти не влияет на нейрон Б. На самом упрощенном уровне все, что происходит, — это синаптическое влияние нейрона А на нейрон Б, повторенное триллионы раз в огромной интерактивной сети. Каждый нейрон получает сигналы от 100 000 других нейронов. И ему нужно подсчитать входные сигналы и принять решение с учетом всего шквала входящей информации, всей возбуждающей и тормозной трескотни, всех “да” и “нет”, сыплющихся в данный конкретный момент, — кто побеждает, “за” или “против”? И если “за”, то нейрон посылает свой собственный сигнал, чтобы воздействовать на более обширную сеть. Задача каждого нейрона — снова и снова принимать это единственное решение. Из кажущегося хаоса повторяющихся простых действий возникают сложные модели.

Нейронные сети, биологические или искусственные, прекрасно обучаются решению сложных задач. К примеру, если вы хотите научить искусственную нейросеть распознавать лица, то давайте ей входящую пиксельную информацию о лице с цифровой камеры, а нейросеть в ответ будет выдавать вам информацию о том, чье оно. Между входом и выходом данных — лабиринт нейронов и синапсов, по которому течет информация. Вначале нейросеть плохо справляется с задачей, связывая лица с конкретными именами наобум. Но с каждой попыткой она получает обучающий сигнал. По мере обучения она вносит изменения в то, какой нейрон куда передает сигнал, насколько крепка каждая из связей и тормозит она или возбуждает. В итоге, оттачивая настройку синаптической схемы, нейросеть научается тому, что должна делать. Увидев лицо Джима — в тени или на свету, улыбающегося или хмурого, — машина свяжет входящую зрительную информацию с нужной исходящей и скажет вам: “Это Джим”. Никто заранее не знает, как должна выглядеть подходящая схема синапсов. Невозможно спроектировать качественное устройство распознавания лиц во всех подробностях. Система обучается методом проб и ошибок, пока не возникнет схема связей, позволяющая успешно справляться с задачей.

Учитывая последние сто лет работы с биологическими нейронами и недавние успехи искусственных нейросетей, большинство нейробиологов сегодня полагают, что суть мозга кроется в схеме связей между его нейронами. С этой точки зрения, если бы мы могли измерить все нейроны в мозге какого-то человека, расписать, какие из них с какими связаны, и описать синапсы между

ними, — мы бы получили сущность этого человека. Описанная гипотетическая карта всех нейронов и их синаптических связей называется “коннектом”[6] — причем слово, похожее на “геном”, выбрано не случайно. И вот почему: если ученым удалось картировать человеческий геном (достижение, которое когда-то считалось невозможным), то они смогут осилить и еще более сложную технологическую задачу — картировать человеческий коннектом. У каждого человека есть свой уникальный коннектом, определяющий уникальный разум.

За последние десять лет произошел некоторый прогресс в его картировании. Опубликован полный коннектом одного из видов круглых червей (*Caenorhabditis elegans*), а позже — коннектом мухи-дрозофилы[7]. Кроме того, ученым удалось взять небольшой фрагмент коры головного мозга мыши, заморозить его, разрезать на невероятно тонкие пластины, отсканировать каждую и воссоздать бóльшую часть нейронов, их переплетающихся пучков и синаптических связей[8]. Этот метод пока не охватывает все синапсы, но вполне возможно, что в ближайшем будущем мы сможем получить полный коннектом нескольких миллиметров мозга мыши.

Национальные институты здравоохранения (NIH) сейчас финансируют проект “Коннектом человека” — грандиозную затею ученых всего мира с конечной целью картировать коннектом человеческого мозга. Для его изучения можно использовать методы МРТ-сканирования, которые в состоянии показать сети нейронных связей со все большей детализацией[9]. Такое неинвазивное сканирование удобнее для испытуемых, поскольку для него не требуется замораживать и резать мозг. Стоит добровольцу

несколько часов полежать в сканере для МРТ — и вуаля, появляется скан удивительно высокого разрешения. (Я много раз проходил сканирование собственного мозга. Это не то чтобы очень приятно, довольно скучно, я часто засыпал, но результат выглядит потрясающе.) Но эти так называемые карты человеческого коннектома не дают достаточно подробного разрешения, чтобы увидеть отдельные нейроны и синапсы. Они показывают структуры большего масштаба — на уровне того, как один участок коры размером с горошину соединяется с другим. Нейробиологи надеются, что по мере совершенствования технологии сканирования можно будет измерять связи в человеческом мозге в лучшем разрешении.

Описанный выше прогресс в искусственных нейронных сетях и измерении коннектома человеческого мозга, казалось бы, позволяет оптимистически смотреть на перспективы оцифровки психики. Мы вроде как знаем, что именно нужно измерить в мозге, а моделировать нейроны уже умеем. И тогда наверняка, учитывая скорость прогресса, перенос личности станет возможным в ближайшие пару десятков лет. Я не разделяю этого оптимизма. Это случится, но не так скоро. Дальше объясню, почему мне кажется, что пока мы очень далеки от него.

Посмотрим еще раз на цифры, которые я приводил выше. Человеческий мозг содержит примерно 86 млрд нейронов и, вероятно, около 100 трлн синапсов — по минимальным прикидкам[10]. Мне еще не встречались устройства, способные отсканировать и измерить 100 трлн чего бы то ни было. Масштабы задачи превосходят возможности современных технологий.

Сканеры для МРТ в настоящее время могут показывать мозг в разрешении приблизительно полмиллиметра, что уже является потрясающим техническим достижением. Но нейроны намного меньше, а синапсы еще мельче. Чтобы увидеть синапс, вам нужно сканировать с разрешением в микрометр — тысячную долю миллиметра. В таком масштабе у вас есть шанс увидеть вздутия на нейронах и предположить, что это синапсы.

Но даже и этого высокого разрешения окажется недостаточно. Ведь нужно будет не только определить, синапс ли вон тот мутный пузырь, но и понять, что это за синапс и какой мощности сигнал может проходить по нему от нейрона к нейрону. Физически более крупные синапсы чаще проводят более мощные сигналы, так что для первичного грубого измерения синаптической мощности потребуется измерить размер синапса. Вам необходимо ясное, подробное изображение каждого синапса в лучшем разрешении, чем микрометры. Ваше сканирующее устройство также должно уметь распознавать, тормозной это синапс или возбуждающий. Для этого, скорее всего, потребуется сканирование на присутствие определенного типа молекул в каждом синапсе. Технологии столь подробного сканирования попросту не существует. Вероятно, при помощи морфологии — знания о различных формах синапсов — удастся выдвигать более или менее достоверные предположения о том, что за синапс перед нами, не прибегая к химическому анализу, но даже такой обходной путь потребует невероятно высокого разрешения, порядка тысячных долей микрометра. Речь идет не об усовершенствовании технологии МРТ. Я говорю об электронном сканирующем микроскопе, используемом для мертвых тканей. А применительно

к живому мозгу — о новых технологиях сканирования, которые еще не изобретены.

В мозге сотни, возможно, тысячи различных видов синапсов[11]. Например, щелевой контакт включает в себя прямую электрическую связь между нейронами. Она чрезвычайно быстра и надежна, а также необходима для нормальной работы некоторых областей мозга, которым требуется мгновенная и точная передача данных. Другой вид синапсов работает как брызгалка, посылающая веером облака химических веществ, которые воздействуют на весь ближайший участок мозга, а не на один конкретный соседний нейрон. В некоторых синапсах содержатся несколько различных химических медиаторов, и в разных обстоятельствах выделяются разные вещества. Ряд синапсов лучше приспособлен к быстрым изменениям (они участвуют в кратковременном обучении), другие же более устойчивы. Всё это разнообразие соединений между нейронами, все эти оттенки и нюансы, включая и те, которые, как можно предположить, еще не открыты, их скорости, мощности и адаптивность — весь этот набор нужно будет считать с человеческого мозга, чтобы создать его коннектом.

Даже если бы нам удалось сосканировать все эти данные от нейронов и синапсов, нам бы пришлось разбираться также и с глиальными клетками, которые часто упускают из виду[12]. Большинство нейробиологов фокусируется на изучении нейронов — отсюда и название профессии. Но в мозге полно и других клеток (их на порядок больше, чем нейронов[21]). Когда-то глиальные клетки считали просто вспомогательными (чем-то вроде скелета, который придает мозгу форму) или же служебными,

которые обеспечивают нужды нейронов и убирают за ними. Но оказалось, что глия обладает свойствами, напрямую связанными с обработкой информации. Некоторые глиальные клетки выделяют химические вещества, воздействующие на нейроны и синапсы. А кое-какие даже испускают такие же электрохимические сигналы, какими сообщаются между собой нейроны. Глиальные клетки плохо изучены в том, что касается их функций, но они не так принципиально отличаются от нейронов, как казалось когда-то. Вот лучшее, что можно сказать сейчас, — наши знания механизмов обработки информации в мозге пестрят огромными белыми пятнами.

Я всего лишь хочу подчеркнуть, что не следует торопиться первыми попробовать новую технологию, как только она появится — ее потребуется еще оттачивать и оттачивать.

Предположим, в будущем изобретут технологию сканирования, при помощи которой окажется возможным видеть в мозге синапсы. Бета-версия, скорее всего, будет сканировать только нейроны, без глии. Вероятно, в силу необходимого упрощения, синапсы удастся отнести только к 100 категориям. Мощность синапсов будет обозначаться приблизительным числовым значением — скажем, одним из 100 возможных на шкале, без дополнительной нюансировки. Сканер, как представляется, не сможет улавливать действие гормонов, которые распространяются по мозгу. У него могут быть невероятные показатели успеха: он верно отметит 99,99% ваших реальных синапсов, но упустит довольно много подробностей, которые важны для общей картины. В результате мы получим как бы эхо оригинала, но нам не удастся предугадать, насколько искореженным и жутким получится разум, который

возникнет из этого набора данных, если на его основе построить искусственный мозг. Вдруг он будет больным, заторможенным, с искаженными эмоциями, неспособным сосредоточиться?

Нарушить нормальное функционирование мозга легко. Даже ничтожные количества некоторых веществ способны вызвать боль, спутанность сознания, галлюцинации и припадки. Сотрясение мозга, которое разрывает ткани и вызывает отек, может привести к месяцам, а то и годам, затуманенного мышления и эмоциональной неустойчивости. Даже крошечные несоответствия иной раз приводят к значительным последствиям. Необходимо, чтобы имитационная версия мозга работала в точности как оригинал, иначе ее опыт окажется поистине кошмарным. Я бы подождал до версии 1000, когда устранят все помехи. Первым подопытным кроликам придется несладко.

Если присмотреться к затруднениям, а именно: колоссальным объемам данных, которые нужно сканировать в мозге, необходимой степени детализации сканирования в доли микрометра, отсутствию на данный момент необходимых нейробиологических знаний, — подмывает попросту отказаться от этого замысла и смириться с тем, что мечта о переносе личности на искусственные носители несбыточна. Да, она несбыточна с сегодняшними технологиями. Даже со всей возможной отладкой сегодняшние технологии остаются на расстоянии световых лет от нее.

Но в то же время я абсолютно уверен, что когда-нибудь это случится. Люди умеют справляться с технологическими задачами. В 1916 г. Эйнштейн предсказал существование гравитационных волн[13]. Но

он считал, что предсказанный эффект будет настолько мал — в 10 000 раз меньше ядра атома, — что его окажется технически невозможно подтвердить. Эйнштейн не мог представить себе машину достаточной чувствительности — ни в каком будущем. Почти ровнехонько через сто лет такую машину построили и подтвердили существование гравитационных волн[14]. Бьюсь об заклад, что новую технологию изобретут, новые возможности появятся и перенос личности станет реальным. Не могу предсказать, когда именно, поскольку это зависит от создания неизвестных и непредставимых на сегодняшний день машин нашими вдохновенными потомками. Если вы хотите временно́й прогноз, то с учетом темпов изобретения принципиально новых технологий сканирования я бы сказал, что не менее века, а то и значительно больше, но я могу оказаться и неправ — вдруг вдохновение посетит кого-то раньше.

И я все же вновь советую не спешить пользоваться первыми моделями.

В переносе психического мира есть две составляющие: первая (как я уже упоминал) — сканирование нужной информации из мозга, который вы хотите скопировать, вторая — создание работающей имитации этого мозга. Предположим, что технические трудности преодолены. Изобретена подходящая машина для сканирования мозга, которая отражает все необходимые подробности его устройства. Теперь нужно при помощи этих данных создать работающую имитацию мозга.

Может показаться, что вторая часть — имитация — задача посложнее, но на самом деле она уже решена. Оборудование для нее готово. Искусственные нейроны и нейронные сети вошли в обиход. Если

нужно будет добавить дополнительные виды синапсов или откалибровать воздействия вроде гормональных, то это не представит принципиальной сложности для имитирования. Известны и понятны даже нейросети, состоящие из миллионов искусственных нейронов. Фирмы по всему миру стремятся создать системы, которые могли бы составить мозгу конкуренцию по сложности. Например, проект “Голубой мозг” при помощи суперкомпьютеров имитирует массивные скопления нейронов, подобных тем, что присутствуют в мозге. Исследовательские группы проекта “Человеческий мозг”, а также Алленовского института, Google Brain, DeepMind, Cogitai и многие другие работают над созданием крупномасштабных систем из сетей искусственных нейронов. Построение сети из 86 млрд нейронов с 100 трлн синапсов все еще превышает возможности сегодняшних технологий. Но они развиваются очень быстро, особенно с наступлением эры квантовых компьютеров. Несомненно, у нас вскоре будет достаточно “огневой мощи”, чтобы сымитировать нейросеть, сравнимую по масштабам с человеческим мозгом.

Столь бурный технологический прогресс — одна из причин сегодняшнего оптимизма в отношении оцифровки психики. Самый заметный и крупный фрагмент загадки практически решен. Но важно не забывать, что искусственная нейронная сеть масштабов человеческого мозга, пусть это и колоссальное достижение, — все еще не то же самое, что оцифрованная личность. Сама по себе сеть, без правильной схемы связей между ее 86 млрд нейронов, — бесполезное скопище цифр. Это как если бы мы разработали принтер, который может напечатать искусственный мозг, и придумали материал, из которого его печатать, но не решили задачу, как

измерить нужные данные в настоящем мозге, чтобы загрузить их в принтер. Без данных мы напечатаем бессмысленное месиво.

Если воспользоваться выражением Дэвида Чалмерса, можно сказать, что у нас есть подлинная трудная проблема переноса личности: сканировать мозг с достаточной степенью детализации.

Представим себе, что каким-то образом мы сумели построить имитацию вашего мозга. Мы его отсканировали и воссоздали нейронные сети. На следующем шаге предполагается вселить копию мозга в тело — без него непонятно, какой опыт будет копиться у вашего искусственного мозга, плавающего в цифровой бочке[15]. Ведь если имитация мозга обладает теми же свойствами, что и мозг живого человека, то в такой ситуации, как мне думается, она может почувствовать дезориентацию и потерять ощущение собственного “я”. Ваше личное “заземление” коренится в вашем теле. Вы знаете, где находитесь физически, где ваши руки, ноги, туловище — это дает вам первичную опору[16]. Если не будет ее, а также контакта с окружающим вас миром, *телесного воплощения,* — подозреваю, что вам придется переживать спутанность сознания, напоминающую наркотический бэд-трип.

Но сейчас нужно решить, какое вам дать искусственное тело: материального робота, который может ходить по реальному миру, или имитацию тела, обитающую в мире виртуальном? Робот ограничен. Мне кажется, селить имитацию мозга в уязвимое, смертное тело — значит упустить перспективы гибкости и долголетия, которые может дать перенос личности на искусственные носители.

Когда моя лаборатория изучала, как мозг управляет движением, мы создали имитацию человеческой руки[17]. Не сделанную из вещества конечность робота — ее нельзя было пожать, она не могла ощупывать предметы. То была виртуальная рука, созданная из данных на компьютере. Все, что мы видели, — матрица чисел на экране. У руки имелось все необходимое благодаря сканам реальной человеческой конечности. Все кости, все связки, все мышцы. У нее была мускульная сила, вязкость, инерция, тяготение. Ее мышцы состояли из отдельных волокон — быстро- и медленносокращающихся. Мы дали руке сенсорные нейроны, альфа-мотонейроны, бета-мотонейроны и гамма-мотонейроны. Сконструировать руку-робот, на столь высоком уровне повторяющую устройство человеческой руки, стоило бы миллионы долларов и годы разработок — и не факт, что у нас бы получилось. Виртуальная же рука обошлась всего в несколько тысяч долларов за компьютер и несколько месяцев работы.

Если нам удалось сымитировать человеческую руку, используя ограниченные вычислительные ресурсы середины первой декады XXI в., то уже, должно быть, становится возможным создание реалистичного виртуального человеческого тела — во всех подробностях его костей, мышц, нервов и кожи. Пока не слышал, чтобы кто-то этим занимался, но надо полагать, что можно создать суперреалистичный аватар в компьютерной игре, похожий на человека изнутри и снаружи.

Когда мы создадим имитацию мозга и виртуальное тело, нам останется встроить эти элементы в виртуальный мир. Можно обратить взор к компьютерным играм с эффектом присутствия, где

имитируются трехмерные миры с их собственной “физикой”. Технологии уже развились практически до нужного для этого уровня.

По-настоящему убедительный, разработанный до мельчайших видимых деталей виртуальный мир — в котором звук реалистично распространяется по виртуальному пространству, ветерок касается виртуальной кожи, возможно, даже запахи и вкусы воздействуют на виртуальные ноздри и языки — пока еще не существует. Но мог бы. Не нужно изобретать принципиально новых технологий. Это вопрос усовершенствования уже существующих виртуальных миров. Ограничение здесь лежит в мощностях обработки информации. Если объединить мощности всех суперкомпьютеров на Земле, то, как мне представляется, их хватило бы, чтобы сымитировать один-единственный человеческий мозг, тело для него и реалистичную трехкомнатную квартиру, в которой искусственное существо бы поселилось. Чтобы создать более обширный мир для большого количества оцифрованных существ, потребуется значительное увеличение вычислительных мощностей.

Понимаю, что картина переноса личности, которую я рисую, — это странная неравномерная смесь. Некоторые технологии уже достаточно развиты. Некоторые — идут быстрым шагом и разовьются в ближайшие лет десять. Компьютерам нужно больше мощности, и вскоре они ее получат. Но отдельные технологии лежат в отдаленном будущем — возможно, придется ждать еще несколько столетий. Части задачи, касающиеся информационных технологий, уже намного ближе к реализации, а те, что связаны с биологией мозга, будут развиваться медленнее. Но я не вижу никаких принципиальных препятствий, которые бы помешали достичь

оцифровки психики. Рано или поздно — скорее поздно — сознания людей оторвутся от биологического мозга и переселятся в искусственный формат.

Перенос личности не видится мне антиутопией. Да, в этом есть значительные риски, но есть и значительный потенциал. Мы, люди, умудряемся хоть и беспорядочно, но разбираться в том, что работает, а что нет, и мне кажется, что перенос психики породит культурный и этический беспорядок, который сам с собой постепенно разберется. Но перед тем как перейти к описанию вероятных преимуществ оцифровки психического мира, я кратко расскажу о пяти возможных подводных камнях.

Подводный камень 1. В сегодняшнем социальном климате, учитывая, как быстро мы отбрасываем устаревшие технологии и каждые несколько лет переходим на устройства поновее, оцифровывать психику было бы не очень практично. Вам повезет, если ваша цифровая версия протянет лет десять, прежде чем перестанет быть совместимой с новой версией операционной системы. Мы избавимся от используемой программы, как избавились от WordStar, ископаемого текстового редактора 1980-х гг., который настолько устарел, что немногие вообще помнят о его существовании. Чтобы платформа для переноса личности имела хоть какой-то смысл для продления человеческой жизни, нам придется пересмотреть свой капиталистический, потребительский подход к информационным технологиям, требующий постоянного обновления.

Подводный камень 2. Объем памяти человеческого мозга огромен, и его предел еще не был достигнут, но теоретически он существует. По мере того как

имитация мозга будет накапливать воспоминания, ее синапсы станут перестраиваться, и в какой-то момент она больше не сможет сохранять новые воспоминания, не повреждая старых. Не думаю, что кто-то знает, какой у имитации будет объем памяти или когда мы исчерпаем ее ресурс. Наверное, речь идет о веках. Возможно, инженерам удастся добавлять дополнительные синапсы в области мозга, связанные с памятью (такие как гиппокамп), чтобы периодически наращивать память сымитированному мозгу. Или же ему придется обходиться движущимся окном памяти, в котором события последних нескольких сотен лет сохраняются с высокой четкостью, а все предшествующие мутнеют.

Подводный камень 3. Какие права будут иметь виртуальные и биологические сознания?[18] Чтобы необходимая технология заработала, имитации чьего-то мозга придется пройти через экзистенциальный ад, перерождаясь снова и снова, пока метод совершенствуется. Если тестовый разум выйдет не совсем таким, как ожидалось, — допустимо ли убить его и начать заново? Если мы сделаем несколько идентичных копий одного человека, означает ли это, что каждой отдельной копией можно пренебречь, поскольку она заменима, — или у них всех есть свои индивидуальные моральные права? И так ли важна изначальная биологическая личность-источник, если в каком-то виде ее психика сохраняется? Другими словами, что происходит со святостью жизни и неприкосновенностью личности после того, как вы уже сняли с себя три копии?

Подводный камень 4. Во многих религиях жизнь после смерти обещается как награда за выполнение правил. И мощью этой награды часто злоупотребляют.

Наградой небес подпитывалось кровавое насилие средневековых крестовых походов. Прельщаются райским вознаграждением и современные террористы-смертники. Но лидеры масс, раздувающие прелести загробной жизни, не могут ничем подтвердить свои слова. Представьте теперь, какой убедительной силой будет обладать объективно достоверное посмертие. Можно взять телефон и написать или позвонить тем, кто уже там, и даже прочитать их отзывы о новых фильмах. В чьих руках окажется контроль над достоверным посмертием — тот и станет править миром. Чуть ли не каждый из кожи будет лезть в надежде на бессмертие, даже если для этого придется совершать этически сомнительные действия — сомнительные даже для абсолютно бессовестных. А уж представить себе новую технологию на службе у недобросовестных политиков — проще простого.

Подводный камень 5. Чей психический мир будем оцифровывать? Богатых? Умных? Политиков, власть имущих? В порядке живой очереди? Ресурсы могут оказаться жестко ограничены, приводя к этически неоднозначной конкуренции. Или в далеком будущем память и скорость обработки информации настолько подешевеют, что это вообще не представит никакой проблемы? Может быть, платформа переноса личности окажется новым YouTube, к которому любой сможет присоединиться? Или на ней станет действовать сетевой нейтралитет? Или же, напротив, у кого-то канал будет шире, чем у других? Надеюсь, что кого бы ни выбрали для оцифровки, — это будут достойные и порядочные люди: ведь с продленным сроком жизни они, скорее всего, смогут оказывать большее влияние на остальных.

Даже если предположить, что все пойдет хорошо и мы избежим основных подводных камней, трудно понять, что будет означать перенос личности на философском или культурном уровне. Задумайтесь над простейшим вопросом: это действительно вы?

Вы не хотите умирать и поэтому в один прекрасный день отправляетесь в клинику оцифровки, пять часов лежите в сканере, вокруг вас мигают лампочки и попискивают сигналы, после процедуры вы ощущаете свои суставы слегка затекшими, но в целом чувствуете себя отлично. Будем великодушны и представим, что технология сработала безупречно. Ее уже протестировали и отладили. С ее помощью все ваши синапсы записали в достаточном разрешении, чтобы воссоздать ваше уникальное сознание. И ему выдается виртуальное тело стандартного образца — достаточно удобное, с вашим лицом и голосом, в комфортной для жизни виртуальной среде. Представим, что все это сбылось.

Кто этот второй вы?

Первый вы — назовем его вы-биологический — заплатил целое состояние за процедуру, но вышел из клиники не менее смертным, чем вошел. Вы остались биологическим существом и когда-нибудь умрете. По дороге домой вы думаете: “Лохотрон какой-то”.

В то же самое время ваша имитация просыпается в виртуальной квартире и чувствует себя в точности как старый добрый вы. Ее опыт непрерывен: она помнит, как пришла в клинику, заплатила карточкой, подписала согласие, легла на стол. Имитация припоминает, что ее усыпили, а затем она проснулась в другом месте. У нее ваши воспоминания, ваш характер, ваш образ мышления и эмоциональные причуды. Она садится в новой постели и говорит:

“Ничего себе, сработало! Не верится, что я здесь. *Определенно* деньги не на ветер выброшены!”

Я не буду больше говорить “имитация”, поскольку этот разум — вариант вас. Назовем его вы-виртуальный. И вот вы-виртуальный решаете освоиться. Вы выходите из квартиры в погожий солнечный день, вокруг вас виртуальная версия Нью-Йорка. Звуки, запахи, виды, люди, ощущение тротуара под ногами — все на месте, только мусора поменьше, и крысы безопасны и здоровы, они здесь лишь для местного колорита. Вы заговариваете с незнакомцами — в настоящем Нью-Йорке вам бы никогда в голову не пришло это делать, чтобы случайно не схлопотать. Здесь у вас не может быть травм: виртуальное тело не ломается. Вы останавливаетесь выпить латте в живописном кафе, присев за кованый столик на улице. Вкус какой-то не такой. И вы не чувствуете, как кофе попадает в желудок. Потому что ничего туда не попадает — это ненастоящая пища, да и желудка у вас нет. До вас доходит, что вам, кажется, больше никогда не понадобится в туалет. На вид текстура столика не очень-то шершава и зерниста, хотя он выглядит немного тронутым ржавчиной. У ваших пальцев нет отпечатков — подушечки гладкие, чтобы на подобных деталях сэкономить память. И дышится как-то не так. Если задержать дыхание, голова не закружится, потому что в виртуальном мире нет кислорода. Вы обнаруживаете при себе виртуальный смартфон и набираете номер, который раньше принадлежал вам, — тот, что, согласно вашему опыту, был у вас всего несколько часов назад, когда вы входили в клинику.

Вы-биологический берете трубку.

— Привет, — говорите вы-виртуальный, — это я. То есть ты. Как дела?

— Хуже некуда, гад ты этакий, вот как. Сижу дома, ем мороженое. Не могу поверить, что повелся и потратил такую кучу денег зря.

— Зря? Чувак, ты не поверишь, что тут творится! Да, текстуры местами подкачали, но я точно найду, чем заняться. Я проходил мимо кинотеатра и книжного, тут есть деньги, поэтому, слава богу, можно ходить по магазинам, и говорят, будто симулятор “Звездных войн” настолько реалистичен, что попадаешь в сам фильм и можешь побыть вуки. А помнишь Кевина, парня, который на прошлой неделе от рака умер? Он тоже здесь! У него все нормально, работает на той же работе. Три раза в неделю связывается по скайпу со своей прежней студией йоги и ведет занятия по фитнесу. Однако его девушка из реального мира ушла от него к другому, который еще не умер, так что ему фиговато. Но тут полно народу, новую найдет…

Придется приложить некоторое усилие, чтобы притормозить, ведь вся эта ситуация может стать бесконечным источником шуток. Но за мелкими подробностями лежит большая философская загвоздка, с которой кому-то придется рано или поздно столкнуться. Как соотноситесь вы-биологический и вы-виртуальный?

Мне больше по душе геометрическое представление ситуации. Представьте, что ваша жизнь — это ствол буквы Y. Вы рождаетесь у ее основания и по мере того, как вы растете, ваша психическая жизнь формируется и меняется по определенной траектории. Затем вы идете на сканирование, и с этого момента Y разветвляется. Теперь есть две траектории, каждая из которых — равно и полноправно вы. Скажем, левая ветвь — это вы-виртуальный, а правая — вы-

биологический. Та часть вас, которая будет жить неопределенно долго, представлена и стволом Y, и левой ветвью. Как ваше детское “я” продолжает свою жизнь в вас-взрослом человеке, так и ствол Y продолжает жить в вас-виртуальном. Когда сканирование окончено, две ветви Y расходятся по разным жизненным путям, накапливают разный опыт. Правая ветвь умрет. Все, что случится с ней *после* точки расхождения, не получит бессмертия — если только вы не решите отсканироваться еще раз, тогда появится новая ветвь и геометрия усложнится.

Вместо единственного вас возникает замысловатая топология — гипер-вы с двумя или более ответвлениями. Одно из них всегда будет смертным, а остальные получат неопределенно долгий срок жизни — в зависимости от того, сколько будет поддерживаться компьютерная платформа.

Вам может показаться, что, раз вы-биологический живете в реальном мире, а вы-виртуальный — в виртуальном, вы вместе никогда не встретитесь, а значит, от вашего одновременного существования не должно возникнуть никаких затруднений. Но разве нужно в наши дни встречаться с людьми лично? Все равно мы общаемся в основном электронно. Виртуальный и биологический вы представляете собой два полностью функциональных, интерактивных, дееспособных экземпляра вас, конкурирующих в одной и той же расширенной взаимосвязанной социоэкономической вселенной. И вы запросто можете встретиться в видеочате или при помощи какой-нибудь еще новой технологии.

Положим, этой экзистенциальной путаницы можно и избежать — если, скажем, хранить данные сканов мозга, не активируя имитацию. Вы будете

периодически захаживать в клинику, чтобы делать резервное копирование мозга. Когда вы умрете (от несчастного случая, болезни или старости), то активируется и превратится в “живую” имитацию самая свежая копия — или та, которая вам полюбилась больше остальных и которую вы отметили в завещании.

Другой, более пессимистичный вариант ограничения количества сосуществующих копий — это специальный закон, который будет, к примеру, предписывать уничтожение вашего биологического тела и переработку его на удобрения, как только ваши психические процессы успешно оцифрованы и ваша имитация активирована. Тогда богатые и власть имущие будут пытаться подкупить блюстителей этого закона, чтобы его нарушить. Они бы желали получить пять или десять копий себя, ну и оригинал тоже сохранить. Потребовался бы, как во второсортном голливудском фильме, спецагент, который бы выслеживал и уничтожал незаконные копии. Вот как все усложняется.

Но может быть, система останется хаотичной и не знающей правил, без ограничений на количество копий каждого человека — и это приведет к революции общественных представлений об идентичности и неповторимости.

На простейшем уровне перенос личности на искусственные носители позволит сохранять людей в бессрочном посмертии. Семьи смогут приглашать виртуальную бабушку по видеосвязи на рождественский ужин, поставив планшет на постамент во главе стола, — если, конечно, она захочет уделить время своей биологической семье, учитывая все ее новые возможности на виртуальном

поприще. Может быть, она подобрала себе мускулистый аватар и штурмует виртуальные горные вершины.

Задумываясь о преимуществах оцифровки психики, люди представляют себе примерно такое вот идеализированное посмертие. Но, в отличие от традиционного рая, это не отдельный мир. Он сплавлен воедино с миром реальным. Как вы взаимодействуете с ним? Если вы ведете типичный для современного человека образ жизни, то контакт с людьми, находящимися физически рядом с вами, занимает меньшую часть вашей жизни. Связь с остальным миром происходит практически исключительно цифровым образом. Вы получаете новости на экран или в наушники. Вы узнаёте о далеких странах из электронной среды, и эти страны обретают для вас реальность. Политики, знаменитости, даже друзья и семья могут существовать для вас только в формате цифровых данных. Люди работают в виртуальных офисах, где знают коллег лишь по видео и сообщениям. Любой из нас мог бы уже сейчас жить в виртуальном мире — CNN, Google, YouTube, Facebook, Twitter и сообщения обеспечивают нам бесперебойный поток информации. Мы живем в странной мультивселенной, каждый в своем виртуальном пузыре, некоторые из них объединяются и затем разделяются в реальном пространстве, но все наши пузыри связаны через глобальную социальную сеть. Если будет создано виртуальное посмертие, то населяющим его людям (с теми же характерами и потребностями, что были у них в реальной жизни) окажется незачем изолироваться от нас. Для них изменится очень немногое. Виртуальный и реальный миры соединятся социально, политически, экономически в одну

расширенную и постоянно растущую цивилизацию. Виртуальный мир запросто сможет стать просто еще одним городом на Земле.

Десятки тысяч лет люди передавали знания из поколения в поколение, рассказывая истории о прошлом. Культ предков наверняка играл в этом важную роль[19]. Постоянно держа в голове, что предки окружают их, наблюдают за ними и незримо их поддерживают, люди ушедших эпох полагались на их мудрость, которая помогала и живущим.

Сомневаюсь, что, когда около 6000 лет назад в древнем Шумере изобрели письменность[20], кто-то понимал, до какой степени это преобразит жизнь нашего биологического вида. Изначально письмо было экономическим инструментом для учета сведений о том, кто что и за сколько продал. Но потом произошли фундаментальные изменения. Письменность позволила людям говорить напрямую через поколения. Резко возросли объем и точность данных, которые можно было накапливать и надолго сохранять. Для науки и технологии особенно важной стала аккумуляция максимально точной информации, записанной учеными-предшественниками. То же можно сказать и об экономической науке, политической теории, религиозной идеологии, течениях в искусстве и любой другой стороне нашего существования. Без письменности современная цивилизация оказалась бы невозможна, поскольку не смогла бы слой за слоем нарастать поверх ушедшего.

Перенос личности на искусственный носитель позволит сохранить больше информации с большей достоверностью. К тому же расширится диапазон доступных к передаче сведений — теперь это не только факты, которые можно вписать в книгу, но и

тонкие нюансы характера, а также умения, которые получается транслировать только при прямом контакте человека с человеком. В чем-то, хоть это и жутковато, мы вернемся к подобию культа предков. Но вместо туманных воспоминаний об их мудрости и подвигах, накопленных в устной традиции, у нас будут собственно предки, с которыми станет возможным поговорить напрямую. Им даже не придется нашептывать свои премудрые советы сквозь пустоту. Они смогут активно участвовать в жизни общества при помощи социальных сетей. На мой взгляд, такое изменение в способах передачи информации и окажется истинной причиной революции, скрытой в новых технологиях. Оцифровка психического мира преобразит наш биологический вид даже сильнее, чем возникновение речи и изобретение письма. Но изменения необязательно будут конструктивными. Как это уже произошло с письменностью, телевидением и интернетом, если общество получает новую технологию, увеличивающую поток информации, то подобная перемена приносит также и приток дезинформации, и вредоносные социальные мемы.

Чтобы снизить риск, мы можем договориться о хранении мудрости только наиболее выдающихся людей. Эйнштейн отличается от обычного человека тем, что у него уникальный набор синаптических весов — удачное сочетание доставшегося ему при рождении и результата обучения в течение всей жизни. Когда вы строите и обучаете нейронные сети, происходит нечто похожее. Вначале у вас есть много похожих сетей с одинаковым количеством нейронов, вы их обучаете и смотрите, что из них выйдет. Некоторые становятся гениями, которые блестяще выполняют поставленную задачу, а некоторые не

оправдывают ожиданий, застревая в субоптимальных состояниях. Сети-“гении” не получаются от добавления дополнительной информации. В конце концов, они состоят из того же количества нейронов. Но им каким-то образом удается выработать удачную схему синаптических связей. А вы даже не можете определить конкретные удачные синапсы. Нельзя сказать: “Вот, смотрите, у этого тормозного синапса — идеальный вес. Какая красота, какой гений!” Нет, инженер не понимает, в чем конкретная причина блестящей работы системы. С помощью обучения и в результате случайностей система обнаружила набор синаптических весов, с которым задачи неведомым образом решаются лучше. Эту сеть и нужно оставить. Будучи неприкрыто прагматичной, оцифровка психики станет сберегать сети, обладающие неуловимым преимуществом. Она будет сохранять людей, в которых выработался хороший набор синаптических связей, дающий полезные навыки. Представьте себе, какую мощь можно обрести, сохраняя эти навыки!

В музыке все пошло бы иначе, проживи Моцарт еще лет двести. Или Бетховен. Да хотя бы Элвис или Джон Леннон. Я не имею в виду, что все изменилось бы к лучшему или к худшему, но подозреваю, что в долгосрочной перспективе музыка изменилась бы меньше. Нам известно из истории, что истинные перевороты в музыке происходили со сменой поколений, когда старая гвардия уходила со сцены, а новая приходила ломать устоявшиеся правила. При успешном переносе личности, вероятно, свежие голоса не пробились бы вперед, а вот старые достигли бы новых высот.

Попробуйте вообразить, как развивался бы английский язык, если бы с нами остались Чосер и

Шекспир — они бы вели беседы, писали, преподавали. Большинству современных носителей английского Шекспир кажется тяжелым для понимания, а Чосер и вовсе невразумительным. Языки меняются со временем, и темпы их изменения хорошо изучены лингвистами. Предположительно, всего лишь 6000 лет назад группа людей говорила на утраченном ныне языке, который называют праиндоевропейским, — и от него отпочковалось огромное разнообразие современных языков[21]. Но лингвистические изменения уменьшились бы, если бы люди, говорившие на том языке, продолжали бы жить среди нас, общались бы с нами, вплетали свои голоса в наш хор. Виртуальные разумные существа могли бы освоить новую манеру говорить — ведь речь меняется со временем у всех нас, — но лингвистические изменения замедлились бы, если бы старшие поколения не умирали. Будь у нас перенос личности, по прошествии 6000 лет мы все еще могли бы говорить на варианте праиндоевропейского.

Не только язык — и мода, и этика, и развлечения, и религия, и культура в целом меняются со сменой поколений. Несмотря на то что у нас есть книги, мы регулярно переизобретаем велосипед. Представьте себе, что люди из прошлого, которые не умирают, придали бы культуре больше инертности. Динамика культурных сдвигов полностью бы изменилась.

Вообразите, как бы это повлияло на политику. Философ и писатель Джордж Сантаяна сказал: “Кто не помнит своего прошлого, обречен пережить его вновь”[22]. Мы так и делаем. Мы повторяем историю, потому что страдаем поколенческой амнезией. Политическая память стирается не полностью: период ее полураспада — примерно два-три поколения. Речь

не о том, что уходит интеллектуальное знание. Школьники все так же учат основные исторические факты, но пропадает непосредственная близость, эмоциональная сила личного воспоминания. Со временем история становится теоретической. После тяжелых уроков Второй мировой войны весь мир остерегался популизма и фашизма. Но участники тех политических событий ушли, в новых поколениях острота воспоминаний утихла, и все больше людей отрицает даже базовые факты. Популизм и фашизм вползают обратно в мир, хотя 50 лет назад это показалось бы невозможным.

Уверен, практически любой без труда приведет примеры циклических исторических ошибок, совершаемых цивилизациями. Неважно, насколько точно сохраняется буквальное знание в наших книгах, — каждое новое поколение, принимая эстафету от предыдущего, жаждет заново изобрести искусство и мудрость жизни. Что было бы, если бы старое поколение не умерло, а так и осталось на вахте? При идеальном раскладе политические умения и мудрость накапливались бы, и мы бы смогли избежать постоянного повторения кошмара. Может быть, политика стала бы больше похожей на науку, наращивая знания и двигаясь вперед. При этом на каждого Ганди, чьей бесконечной жизни мы бы только радовались, нашелся бы Нерон или Гитлер, которого оказалось бы намного труднее уничтожить.

Вообразите теперь, как изменилась бы университетская система. Преподаватель блестяще работает на протяжении 30 лет, учит студентов, заседает в комиссиях — и получает звание эмерита[22]. Он остается в должности и продолжает работать. Потом умирает. Из эмерита его повышают в мортем. Ему больше не нужен кабинет — только

пространство на платформе, куда перенесена его личность. Оттуда он продолжает преподавать и заседать в комиссиях по видеосвязи — в общем, как и раньше, делится своими уникальными знаниями и мудростью. Если это преподаватель истории, он будет особенно ценен для новых поколений, которые захотят из первых рук узнать о том, каким был мир века назад. Только представьте, какие исторические знания оказались бы нам доступны, если бы государственные служащие Древнего Египта могли сами поведать о своей жизни. Но также представьте, как трудно было бы молодым ученым нашего времени получить постоянный контракт в университете, забитом поколениями умерших профессоров.

Обычно глубоких стариков трудно представить активными участниками жизни общества. Их немного, они так хрупки и так сильно отстали от последних тенденций. Я встречал людей — правда, их все меньше — у которых никак не укладывается в голове, что такое компьютер, и которые не знакомы с интернетом. Некоторые даже смартфонами не пользуются, потому что не хотят разбираться в них, — слишком много нового для мозга. Но все остальные живут в эпицентре бума технических изобретений, который выталкивает старшее поколение на обочину. Мы существуем в истончающемся срезе настоящего, в перегрузке информации сегодняшнего дня. Но, хорошо это или плохо, перенос личности на искусственный носитель кардинально все изменит. Старики перестанут быть вымирающим видом. Они будут жить бок о бок со всеми остальными и хранить верность старому, привычному образу жизни. Культурные новшества обычно приходят от младших поколений, а старшие им сопротивляются. Если старшее поколение не сойдет со сцены, оно изменит

демографическую картину и наверняка склонит культуру в сторону большей устойчивости. Вот парадокс — цифровое посмертие, с одной стороны, радикально изменит общество, а с другой — сделает его консервативнее и замедлит его изменения.

Старшие поколения не только останутся жить, они еще и обретут власть. Нет причин полагать, что у живых людей будут какие-либо политические, экономические или интеллектуальные преимущества перед виртуальными. Задумайтесь над тем, кем они работают в нашем мире. От многих требуется физический труд, но как раз эту работу смогут выполнять роботы. Таксисты? Общественные беспилотные автомобили почти уже здесь. Дворники? Кассиры? Строители? Летчики? Скорее всего, подобные профессии через какое-то время попадут под раздачу. Их освоят роботы и искусственный интеллект. Остальная наша работа, наш вклад в мир делается умом, а если его можно перенести на другую платформу, то он окажется в состоянии продолжать трудиться и после смерти мозга. Политик сумеет работать из киберпространства не хуже, чем из физического. То же самое с учителями, психотерапевтами, журналистами, писателями и специалистами по обработке рекламаций. Генеральный директор компании вроде Стива Джобса, у которого сформировался великолепный набор нейронных связей в мозге, позволявший ему достигать потрясающих успехов в работе, сможет управлять бизнесом без физического присутствия. Когда понадобится пожимать руки, у него будет возможность ненадолго арендовать человекоподобного робота (воспользоваться чем-то вроде бот-шеринга) и провести несколько часов в реальном мире, общаясь с людьми вживую. По мне,

так само слово “реальный” в отношении мира отдает предрассудками. Оба мира станут одинаково реальными. Может быть, лучше называть их “мир-основание” и “мир-облако”.

Мир-основание будет полон молодежи — людей младше восьмидесяти, которые еще накапливают ценный опыт. Подразумевается, что их главные задачи — расти, настраиваться, обретать мудрость и квалификацию, прежде чем пополнить ряды мира-облака. Оптимальное сочетание культуры и власти быстро переместится в этот облачный мир. А как может быть иначе, если там соберутся знания, опыт и политические связи? При таком раскладе мир-основание станет своего рода личиночной стадией для незрелых человеческих умов, а настоящая жизнь будет начинаться в мире-облаке.

Давайте поговорим о путешествиях в космосе. Первое оптимистичное, распространенное в культуре представление о будущем дала нам уже упоминавшаяся здесь киноэпопея “Звездный путь”: человечество преодолеет болезни и войны и расселится по Галактике — мирный, увлеченный научными исследованиями биологический вид, путешествующий на кораблях с гиперпространственными двигателями. В космосе нам будут встречаться разнообразные зловредные инопланетяне, которых нам удастся без применения насилия разносить в клочья. В такой вариант будущего вы верите или в какой-нибудь другой из меню научной фантастики — все их объединяет культурное допущение, что наше будущее лежит в космосе и что наш вид колонизирует Галактику.

Но путешествовать в космосе сложнее, чем представляется. Человеческое тело не предназначено

для контакта с токсичной средой открытого космоса. Даже если воздух в нашем корабле будет насыщен кислородом, смертельный дождь космических лучей проникнет почти сквозь любую обшивку. Международная космическая станция, которая сейчас на орбите, защищена магнитным полем Земли, отражающим космические лучи. Чтобы добраться до Марса, людям придется выйти за пределы защитного воздействия нашей планеты, и это может оказаться смертельно опасным. Если лететь дальше Марса, опасность будет возрастать в геометрической прогрессии.

Даже если нам удастся разрешить все технические трудности выживания людей в космосе, останется проблема длительности путешествия. Ближайшая звездная система — Альфа Центавра — находится на расстоянии более четырех световых лет. При полете со скоростью 0,1 *c*[23] (а это намного быстрее, чем могут позволить современные космические технологии), путь в один конец займет 45 лет. Если бы мы изобрели волшебный космический корабль, который в состоянии развивать скорость в 99,9% скорости света, нам бы потребовалось более 100 000 лет, чтобы пересечь Галактику. Человеческая жизнь слишком коротка для осмысленного освоения космоса. Мы могли бы путешествовать в космосе сменяющимися поколениями — возможно, так и будет. Но тогда еще бóльших усилий потребует поддержание человеческой жизни.

Мечта о победоносном расселении нашего вида по всей Галактике, скорее всего, так и не сбудется. Подобное представление о грядущем — не более чем воздушные замки. Наша научная фантастика ошиблась.

Но у проблемы освоения космоса есть решение. Цифровая психика не нуждается в кислороде, подходящем атмосферном давлении или органической пище. Ей нет необходимости везти с собой в герметичной кабине окружающую среду, подобную земной. Кроме того, она может жить неопределенно долго и ускорять или замедлять скорость приобретения сознательного опыта. Представьте себе группу людей — сотни, тысячи или даже это будут двойники вообще всех сознаний, которые когда-либо были оцифрованы. Они существуют на виртуальной платформе в космическом корабле. Задача группы — исследовать. Там никому не тесно, поскольку вся жизнь протекает в виртуальном мире. Члены группы могут удобно рассредоточиться по целому городу либо виртуальной копии Земли или же по любому другому виртуальному ландшафту. Если им захочется, они смогут создать виртуальный звездолет “Энтерпрайз” и жить в нем, как в кино: большие экраны на капитанском мостике будут указывать путь, а маленькие лампочки — забавно перемигиваться. Ученые и техники группы путешественников займутся наблюдением за Вселенной (сквозь иллюминаторы и с помощью приборов), а также корректировкой траектории. В случае скучного многовекового перелета от одной звездной системы к другой можно будет просто повернуть специальную ручку и замедлить скорость обработки информации так, что покажется, будто прошло всего полчаса — как раз подходящий интервал времени, чтобы подстроить траекторию для приближения.

Если, достигнув чужой звездной системы, первопроходцы захотят обследовать одну из планет, им не понадобится пригодная для дыхания атмосфера.

Все, что им нужно, — это чтобы их оборудование не расплавилось от высокой температуры, а гравитация его бы не расплющила. Будь вы членом такой команды, вместо того чтобы надеть скафандр и вылезти из капсулы, вы перенесете свою нейросеть на посадочный модуль и пойдете гулять по новой планете, фотографировать и собирать образцы. А нагулявшись, вернетесь на виртуальную Землю или в любую другую среду, имитация которой ждет вас на базе.

Подозреваю, что прямо сейчас большинство людей предпочло бы, чтобы космос осваивали реальные люди. Как будто космическая цивилизация получится не совсем полноценной, если она ограничена виртуальными умами и роботами для сбора образцов. Мы хотим, чтобы кто-то вроде Нила Армстронга полетел и увидел планету своими глазами. Но мне кажется, что единственная сущностная составляющая нас, т.е. именно то, что определяет человека, — это психика, и когда-нибудь технологии позволят человеческой психике свободно путешествовать по Вселенной.

Мне представляется неизбежной последовательность трех этапов. Во-первых, изобретут способ оцифровывать психику. Психологическая мотивация — желание сохранить разум после смерти тела — так велика, что инженеры и ученые уже работают в этом направлении. Во-вторых, когда будет реализован перенос личности, очевидным применением для него станет возможность посылать человеческую психику на искусственных носителях туда, куда человеческое тело так просто попасть не может. В итоге, это в-третьих, разовьется полноценная цивилизация покорителей космоса, которая сможет за огромные промежутки времени преодолевать

необъятные расстояния и распределяться по всей Галактике. Ключ к такому будущему — осознать, что человеческая психика есть информация, которую можно перенести с физического носителя (мозга) на искусственный. Переломным в развитии космической цивилизации станет момент, когда мы разберемся в том, что такое сознание с инженерной точки зрения.

Хочу еще кое-что добавить о переносе психики, это касается человеческой социальности. Людей тянет друг к другу. Необходимость связи с другими заложена в нашей природе. Да, у нас есть свои отшельники, одиночки и агорафобы, но большинство общается с кучей народу. Нас перестал удовлетворять обмен бумажными письмами, поэтому мы изобрели телефон, а потом электронную почту и мессенджеры. Мы разгуливаем по улицам, не выпуская смартфонов из рук, постоянно таким образом подпитываясь от социальных сетей. Если бы у нас был дар телепатии, то какая-то часть населения ухватилась бы за такую возможность и оставалась бы на непрерывной — и безо всяких посредников — связи с миллиардом желающих.

Оставив в стороне магию и лженаучные заявления, признаемся, что людям еще не удавалось напрямую делиться друг с другом мыслями — главным образом потому, что мы рождаемся без USB-портов в головах. Но виртуальный разум будет состоять из информации, которая обрабатывается и управляется на искусственной платформе. Технически один виртуальный разум вполне мог бы напрямую соединиться с другим.

Не знаю, как именно работала бы эта технология. Красота и простота переноса личности заключается в том, что инженерам совсем не обязательно понимать,

как работает мозг: все, что от них требуется, — снять с него копию. Если скопировать мозг с достаточной точностью, он должен будет работать так же, как оригинал. Но прямое подсоединение одной психики к другой требует лучшего понимания того, как порождаются мысли. Какая информация используется на выходе, какая на входе? Подготовившись к задаче недостаточно тщательно, мы возьмем лишь случайные сигналы психики от одной личности и отправим их личности другого в виде бессмысленного шума. Учитывая сегодняшнюю ограниченность наших знаний в нейробиологии, я даже не могу себе представить, как создать такую технологию. Объединение двух оцифрованных психических миров — задача для очень отдаленного будущего, когда о человеческом мозге станет известно намного больше. Но если это все же когда-нибудь случится, то произойдут фундаментальные изменения в динамике информационных потоков. Вместо сборища виртуальных личностей оцифрованные психики превратятся в центральный узел интеллекта, в котором человеческая индивидуальность окажется утраченной. А те, кто пока населяет биологический мир, станут ментально ущербными, всего-навсего расходуемыми частичками мозга, которые трудятся, накапливая опыт и уповая на возможность умереть и слиться с коллективом.

Подобные картины будущего обычно пугают людей — мне тоже от них страшновато, но, думаю, отторжение — всего лишь страх неведомого. Возможное будущее с переносом психики настолько чуждо, что наши эмоции и мысли буксуют при попытках его осознать.

Мне нравится воображать небольшую группу кроманьонцев у вечернего костра примерно 30 000

лет назад. Один из них рассказывает историю о невероятном будущем. Горы сравняют, леса вырубят, плодородные земли покроют искусственным камнем, а поверх настроят огромных угловатых пещер, которые закроют небо. Величественные животные, на которых мы сейчас охотимся и которых рисуем на стенах наших пещер, — зубры, мамонты, пещерные медведи — вымрут. Возвышенный образ жизни охотника станет невозможен. Мало кто сможет изготовить копье или стрелу, да хоть вообще что-нибудь собственными руками. Большинство людей будет целыми днями сидеть в помещениях с тусклым светом, щелкая маленькими квадратиками на прямоугольной доске, их тела станут мягкими и медленными. Воздух окажется загрязнен произведенными человеком выхлопами и наполнится произведенным человеком же постоянным звоном и грохотом. Я так и вижу ужас и неверие на лицах слушателей, отказывающихся признавать эту антиутопию за будущее.

Но вот она вокруг нас, а мы даже не особенно и возражаем. Немногие захотят вернуться к первобытному укладу жизни, к охоте и собирательству. Мы знаем, что у нашего мира есть изъяны, но не хотим слишком уж менять его. Когда мы представляем себе утопическое грядущее, оно всегда более-менее похоже на наш мир, только без худших его проблем и с парочкой новых технических удобств. Представление о будущем, которое слишком сильно отличается от настоящего, по определению покажется антиутопией. Я думаю, это верно для представителей любой эпохи в истории человечества.

Но, одобряй не одобряй, — мир будет меняться. Искусственный интеллект и перенос личности перестроят его. Люди, населяющие то будущее, скорее

всего, с ним справятся, так же как мы — с сегодняшним миром. Они привыкнут и, наверное, не захотят отмотать время назад, чтобы вернуться в нашу ограниченную и беспорядочную жизнь.

Стараясь заглянуть как можно дальше вперед, самое важное изменение, которое я вижу, — поистине переломное в истории нашего биологического вида — это момент, когда люди поймут сознание. Разобравшись в нем с практической, инженерной точки зрения, мы увидим возможности для потрясающего будущего. В этом будущем психика станет чем-то драгоценным, ее начнут взращивать, лелеять и хранить, ее удастся изъять из исходного биологического носителя и перенести, скопировать, разветвить, ее можно будет поддерживать бесконечно долго и даже, возможно, объединять с другими психиками. Я вижу, как разум отделяется от биологической основы, а грань между человеческим и искусственным интеллектом размывается до невидимости. И я вижу это трансцендентное свойство разума, распространяющееся в космосе и осваивающее Галактику — теоретически, через миллионы лет. Технология психики может стать нашим наилучшим путем в далекое будущее.

Приложение

Как сконструировать сознательное восприятие зрительной информации

Проведем мысленный эксперимент. Наша цель — продумать, как построить машину, которая, увидев яблоко, сможет осознать, что́ она видит. У этого упражнения два возможных практических применения. Во-первых, оно окажется чем-то вроде учебника теории схемы внимания и, в частности, поможет в простейшей форме описать логические основания теории. Во-вторых, это упражнение укажет инженерам общий план действий в направлении искусственного сознания — на что я очень надеюсь.

Я часто привожу яблоко в пример: оно обладает понятными видимыми чертами — простой формой и ярко выраженным цветом. Оснастим нашу машину глазом-камерой и мозгом-компьютером и попробуем узнать, что такого нужно встроить в ее искусственный мозг, чтобы он заявил о наличии субъективного осознанного опыта яблока. У мысленного эксперимента есть ограничение: все, что мы добавляем к гипотетической машине, должно быть доступно для постройки. Некоторые элементы, вероятно, уже имеются в продаже. Другие еще не разработаны, но если они выглядят правдоподобно при современном состоянии технологий, то можно включить и их.

Я представлю три различных версии машины, каждая будет чуть более продвинутой, чем предыдущая, а при описании последней версии приведу доводы о том, что машина обладает, в сущности, таким же сознанием, как человек. Я не имею в виду, что машина живет богатой внутренней жизнью. У нее не будет осознания себя, не будет эмоций, воображения или целей. Она сможет осознавать лишь одно: яблоко. Но даже этот крошечный шаг сам по себе укажет направление работы для создания сознания с расширенным содержанием.

На рис. А.1 наш робот смотрит на яблоко. Ячейка над головой робота отображает информацию, имеющуюся в устройстве, и на этой первой схеме в ячейке содержится только один элемент: зрительная информация о яблоке.

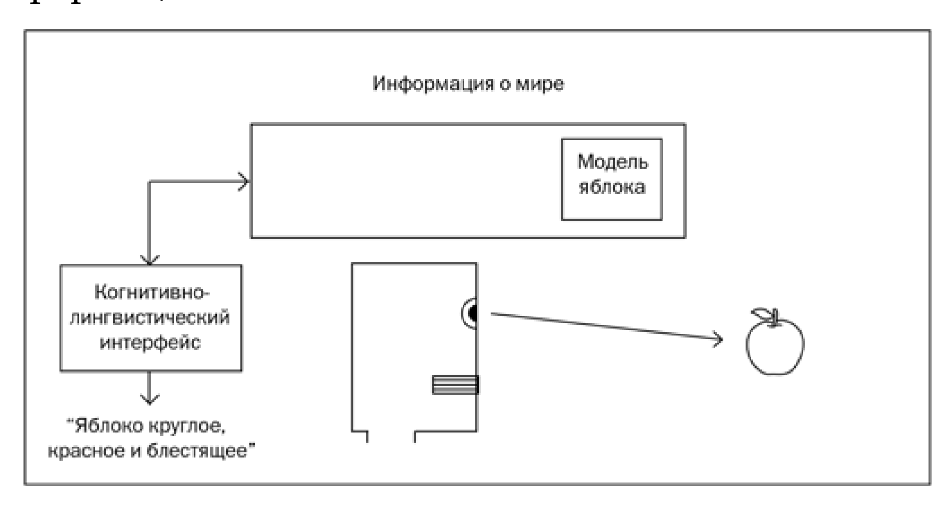


Рис. А.1 Робот получает зрительную информацию через глаз-камеру и строит внутреннюю модель, или пакет информации, о яблоке. Поисковая система (когнитивно-лингвистический интерфейс) позволяет системе отвечать на вопросы, основываясь на внутренней информации

Зрительная обработка данных — одна из самых изучаемых областей нейробиологии. На сегодняшний день зрительная система человека в общих чертах обрисована, хотя и остается множество невыясненных загадочных подробностей. Все, что нам требуется

знать для проекта “создай мозг”, — это что уже были прецеденты постройки искусственных зрительных систем, пусть и ограниченных и упрощенных по сравнению с той, которая действует в человеческом мозге[1]. Давайте наделим нашего робота способностью вобрать зрительный образ через глаз-камеру и составить обширный, подробный комплекс информации о яблоке. Искусственная зрительная система сводит воедино данные о цвете, форме, размере и расположении яблока, создавая нечто вроде досье — комплекса информации, который постоянно меняется по мере поступления новых сигналов. Этот пакет данных иногда называют внутренней моделью. Можно представить ее себе как имитацию яблока.

В реальном, биологическом мозге эти имитации не очень достоверны. Если бы мозг строил подробные, научно точные описания яблок, это было бы пустой тратой энергии и вычислительных ресурсов. Скажем, цвет яблока отчасти достраивается мозгом. В действительности яблоко обладает не цветом, а спектром отражения. Глаза и мозг упрощают этот спектр и присваивают яблоку цвет. Но это, скорее, карикатура на цвет — для простоты и быстроты обработки. Для выживания важна оперативность. Вы явно не захотите заниматься настолько тщательным моделированием, что не сумеете вовремя среагировать на мир.

Все, что я описал выше, не нарушает ограничения, упомянутого в начале раздела. Раз у нас есть камера и компьютер, нам удастся построить систему, изображенную на рис. А.1. Но сможет ли наш робот *осознать* яблоко?

В каком-то смысле, да. Как я уже говорил, исследователи сознания иногда пользуются термином

“объективное осознание” для обозначения того, что информация попала внутрь и обрабатывается[2]. Да, машина с рис. A.1 объективно осознает яблоко. В ней есть эта информация.

Но осознает ли она его *субъективно*? Есть ли у нее субъективное переживание круглости, красноты и блеска — такое, какое было бы у вас, рассматривай яблоко вы? Некоторые ученые будут утверждать, что да — сознание это ощущение, возникающее при обработке информации[3]. Поскольку машина обрабатывает яблоко, она по определению осознает его субъективно. Я называю такие теории “сопряженными с информацией”: сознание — неизбежный побочный эффект обработки данных подобно тому, как тепло — побочный эффект электрических сетей. Исходя из этой точки зрения, если построить вычислительную машину, то обязательно возникнет сознание. Если это так, у нас все готово. Мы успешно спроектировали сознающую машину, хотя сознание так и осталось необъясненным.

Но мне не кажется, что машина готова. Чтобы доказать это, давайте спросим самого робота. Раз уже мы проводим мысленный эксперимент по конструированию мозга, то кто нам мешает встроить в него и лингвистический интерфейс, поисковик вроде Сири (на рис. A.1 он обозначен как когнитивно-лингвистический интерфейс). Он получает вопросы, ищет ответы по доступной внутри машины базе данных и сообщает результат поиска. Поскольку такой интерфейс вполне реально создать на основе сегодняшних технологий, нам “позволено” включить и его в нашу машину, хотя это новое для него применение.

Мы спрашиваем машину: “Что это?”

Машина: “Яблоко”.

Мы: “Какие у яблока свойства?”

Машина: “Оно круглое, красное, блестящее, сверху вмятина, лежит там-то”.

Робот может ответить на эти простые вопросы, поскольку содержит нужную информацию. Более того, его ответы насыщенны и гибки, потому что его знание о яблоке основано на модели, а модель, к которой он обращается, — обширное, если не полное, описание яблока. И вот у нас есть машина, которая в состоянии взглянуть на яблоко, обработать данные о нем и сделать конкретные касающиеся его утверждения. Такую машину еще вполне возможно построить в сегодняшних реалиях.

И тем не менее в качестве теории сознания рис. А.1 чего-то не хватает. Чтобы доказать это, зададим машине очевидный последний вопрос: “А ты осознаешь это яблоко?”

Поисковая система обращается к внутренней модели, но не находит ответа. Она находит кучу информации о яблоке и никаких данных об осознании — что это такое и есть ли у машины это загадочное свойство. У машины попросту нет данных о себе самой. Мы ведь ее спросили: “Осознаешь ли *ты* яблоко?”, а ей не хватает информации о том, что такое это “ты”. Наш вопрос не имеет для машины смысла. В лучшем случае она ответит: “Ошибка вычислений”. С тем же успехом мы можем спрашивать цифровую камеру, осознает ли она фотографию, которую только что сделала. Ничто из того, что мы уже вложили в машину, не приведет ее к утверждению о наличии сознания.

Я принимаю это как постулат базовой логики: машина для обработки данных не может сделать утверждение — не может выдать информацию, — если в ней не содержится информация, которую она утверждает. Машина на рис. А.1 не содержит информации о том, что такое сознание, и уж подавно о том, что она что-то осознает. Нельзя наивно надеяться, что, вложив в машину сложность, мы получим от нее утверждение о сознании. Имея лишь те компоненты, из которых мы ее построили, машина не может содержать информацию о, скажем, зебрах, и поэтому не способна взять и заговорить о зебрах. С инженерной точки зрения лелеять надежду, что машина проснется и заявит: “Я сознаю” — это нарушение логики. Нам придется вложить в машину больше данных, чтобы расширить диапазон возможных ответов.

На рис. А.2 изображена следующая версия нашего проекта “создай мозг”. Сюда я добавил вторую внутреннюю модель: модель себя. Она тоже состоит из данных, которые собираются в мозге. Эта информация может описывать физическую форму и структуру тела — уже упоминавшуюся схему тела[4]. Она способна включать в себя и автобиографические воспоминания, которые играют важную роль в любой психологической модели себя[5]. Нам известно, что человеческий мозг содержит что-то вроде модели себя — скорее всего, очень сложной и распределенной по многим его областям. Мы имеем право дать нашему роботу модель его самого, поскольку техническая возможность предоставить ему некий произвольный комплект информации вполне доступна.

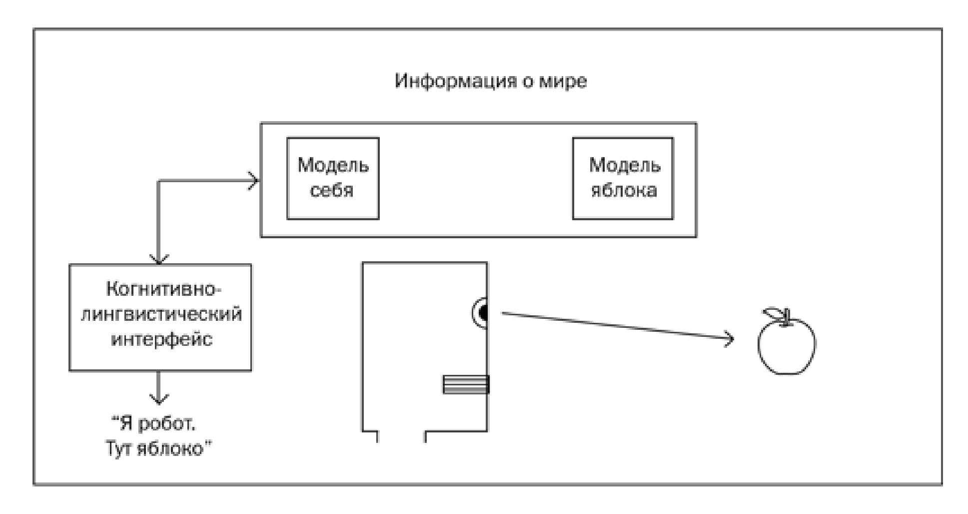


Рис. А.2 Обновленная версия робота со второй внутренней моделью — пакетом данных, описывающих его самого

Теперь спросим улучшенную машину: “Кто ты? Расскажи о себе”.

Со своей новой внутренней моделью робот располагает необходимой для ответа информацией. Он скажет: “Я металлическая личность ростом 1 м 80 см, мои руки сгибаются в локтях и плечах, я родился в лаборатории, пять дней назад у меня был день рождения…” Поисковая система умеет обращаться к информации во внутренней модели себя и выдавать ее наружу. Некоторые данные могут касаться самой машины, а другие — описывать события в прошлом, которые помогают определить ее.

Появилось ли у этого робота сознание? Кто-то скажет, что да. Во многих теориях сознание приравнивается к знанию о себе. Некоторые ученые, разделяющие подобные взгляды, сосредоточивают внимание на схеме тела. С их точки зрения, базовое, примитивное сознание — это знание о физическом “я” и его движениях[6]. Я знаю, что мои части тела принадлежат мне, они часть меня и принципиально отличаются от других предметов вокруг меня. Я знаю свое расположение в пространстве: я существую здесь. Я знаю свою точку зрения: я вижу, слышу и понимаю мир со своей физической колокольни. Другие

исследователи обращаются к абстрактному, психологическому знанию о себе[7]. Исходя из их позиции, я обладаю сознанием, поскольку знаю свой жизненный путь, знаю свои интересы и мотивы. Я могу составить последовательное изложение, объясняющее, кто я такой и почему делаю то, что делаю.

Машина, изображенная на рис. А.2, многое знает о себе. Технически ничто не препятствует вложить в нее сколько угодно этого знания, и я, бесспорно, соглашусь, что оно важно для сознания. Но опять же мне кажется, что и эта схема неполноценна. Чтобы доказать это, спросим машину: “А каковы мысленные взаимоотношения между тобой и яблоком?”

И тут машина залипнет. Поисковая система отправит запросы обеим внутренним моделям, получит гору информации о машинном “я”, отдельно от этого — не меньшую гору о яблоке, но никаких данных не получит о мысленных взаимоотношениях между ними и даже о том, что такое эти мысленные взаимоотношения. Поисковик в принципе не сумеет понять вопрос. С той оснасткой, которую мы ему выдали, робот на этот вопрос ответить не в состоянии. Он не может утверждать, что осознает яблоко, сознание к нему вообще не применимо.

Преимущество внутренних моделей заключается в том, что они отслеживают и предсказывают важные явления в мире. Мир нашего робота содержит два очевидных явления: яблоко и его самогó. Вот мы и дали роботу две внутренних модели: модель яблока и модель себя. Но мы упустили третье, неочевидное явление из его мира: вычислительные взаимоотношения робота и яблока. Чтобы у робота было полноценное описание его мира, нужно

встроить в него третью внутреннюю модель — модель его собственных процессов зрительного внимания.

Человеку зрительное внимание необходимо, поскольку перед ним, как правило, оказывается слишком много предметов — это могут быть яблоко, тарелка, стул, стол, сотни других: невозможно обработать их все с достаточной глубиной. Мозг вынужден расставлять приоритеты и в одно время сосредоточивать ресурсы только на небольшой выборке. В этот конкретный момент лишь несколько предметов побеждают в соревновании: их внутренние модели орали достаточно громко, чтобы повлиять на системы по всему мозгу[8]. Объекты, оставшиеся без внимания, по сути, не обрабатываются — мы даже не замечаем их присутствия, зато то, на что внимание было обращено, обрабатывается как следует. Мозг умеет извлекать подробности и смыслы из этого избранного и решать, как ему реагировать.

Иногда люди пользуются словом “внимание” в разговорном смысле, имея в виду единственный предмет, на котором сосредоточен их интерес. Например, мой интерес может быть сосредоточен на разговоре, которому я посвящаю основную часть мысленных ресурсов, а яблоко, которое я одновременно с этим ем, не имеет для меня особого значения. Но этот термин в научном смысле более широк. Исходя из него, в приведенном примере внимание уделяется и яблоку, и разговору. Их репрезентации в мозге получают достаточно поддержки, чтобы подвергнуться глубокой обработке. В это же время мы не обращаем внимания на сотни других предметов, мыслей и сенсорных сигналов — мы даже не знаем об их существовании. Когда я говорю о внимании, я имею в виду более широкие

взаимоотношения мозга и предметов, на которых он сосредоточивает свои ресурсы.

Зрительное внимание уже конструировалось искусственно, хотя получившиеся искусственные системы оказались намного проще своего биологического, человеческого прототипа[9]. Поскольку машина со зрительным вниманием — это уже существующая и доступная технология, мы можем включить и эту функцию в проект “создай мозг”.

На рис. А.3 изображена третья и последняя версия робота. Здесь он направляет зрительное внимание на яблоко. Прошу вас заметить, что теперь у него три внутренние модели, обеспечивающие его полным описанием трех основных составляющих его мира. Первая модель относится к яблоку, вторая описывает его самого, а третья — тот самый незримый элемент, взаимоотношения внимания между “я” и яблоком. Эта модель внимания обозначается как “схема внимания” по аналогии со схемой тела, описывающей тело.

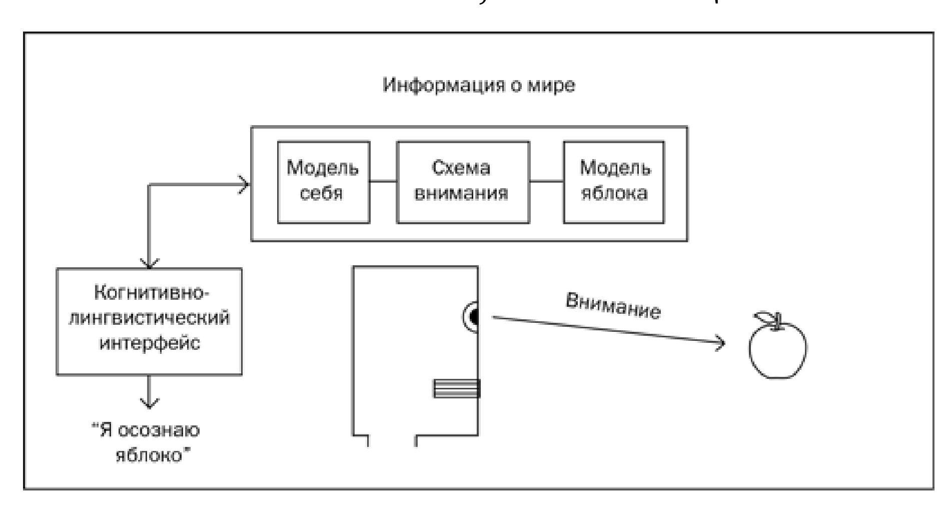


Рис. А.3 Основные элементы теории схемы внимания. Робот содержит модель яблока, модель себя самого и модель взаимоотношений внимания между “я” и яблоком

У нас есть довольно точное представление о том, какие данные должны быть в модели яблока и в модели себя. Но каким данным следует содержаться в

модели внимания? Как она должна описать внимание, чтобы это можно было использовать?

Ее описание не может быть примитивным, на уровне: “Яблоко — это то, чему я уделяю внимание”, иначе модель окажется пустой, не более чем пунктирной линией между двумя другими моделями. Схеме внимания нужна информация о самом внимании — о том, что такое сосредоточивать его. Но, как и упрощенная внутренняя модель яблока, хорошая схема внимания не будет зря тратить ресурсы на данные о неважных для нее технических деталях. Для внимания настоящего мозга эти технические мелочи включают в себя нейроны, синапсы и соревнование между электрохимическими сигналами. В случае робота это провода, чипы и кремниевые логические вентили. То есть составные части механизма, о которых системе знать необязательно. Нашей машине нужно поверхностное, лишенное деталей, описание внимания, уделяемого ею яблоку. Она может описывать его как мысленное удерживание предмета, как способ захвата информации, познания ее и глубокого ее понимания. Или же как нечто, происходящее где-то внутри. Может она описать и некоторые предсказуемые последствия: внимание предоставляет способность реагировать, запоминать, принимать решения.

Теперь спросим машину, изображенную на рис. А.3: “Каковы мысленные взаимоотношения между тобой и яблоком?”

Поисковая система обратится к внутренним моделям и ответит на основе имеющейся информации. Она скажет: “Я мысленно удерживаю яблоко”.

Многообещающий ответ, но мы можем копнуть еще глубже. “Расскажи нам еще об этом мысленном удерживании. Какие у него физические свойства?” Чтобы удостовериться, что машина понимает вопрос, уточним: “Ты знаешь, что такое физические свойства?”

У машины есть схема тела, которая описывает ее физическое тело, и есть внутренняя модель яблока, которая описывает физический объект. Так что машина может сказать: “Да, я знаю, что такое физические свойства. Плотность, вес, импульс, прозрачность — это свойства предметов”. Но, основываясь на доступных ей данных, она возможно, поведает нам (при достаточном словарном запасе): “Мое мысленное удерживание яблока, мысленное удерживание *как таковое* — не имеет свойств, которые можно описать физически. Оно существует, но у него нет плотности, веса, импульса или прозрачности. Удерживание не существует в этих измерениях. Это нефизическая сущность. В каком-то смысле — *метафизическая*. Однако оно действительно имеет физическое местоположение — существует где-то внутри меня. Это часть меня — метафизическая часть, которая позволяет мне вбирать что-то умом. Я не просто обрабатываю информацию, что это яблоко красное: у меня есть мысленное *удерживание* красноты”.

Добавив схему внимания, мы создали машину, которая утверждает о наличии у нее сознания, подозрительно похожего на то, о наличии которого утверждает человек.

Мы сами построили робота и знаем, как он работает, поэтому все еще смотрим на это скептически. Устроим ему дополнительную проверку:

“Спасибо за объяснение, но ты все-таки только машина — и *конечно*, будешь извергать подобную информацию. Ты просто обращаешься к внутренней модели, своей схеме внимания, и выдаешь информацию из нее. Мы, конечно, впечатлены, что твои основанные на модели знания позволяют тебе давать такие гибкие и богатые ответы на вопросы. Но ты же всего-навсего передаешь информацию, которая содержится в тебе”.

Услышав нашу претензию, поисковая система вновь обратится к внутренним моделям и не найдет ничего, подходящего под только что сделанное описание. Вспомните, эти внутренние модели слишком упрощены, чтобы четко сообщить нам, как в действительности работает машина. Поскольку ей доступна лишь ограниченная информация, то она отвечает: “Я ничего не знаю о внутренних моделях, информации или ее обработке. Вот что я знаю: яблоко здесь, я здесь и я владею осознанием яблока. Осознание само по себе — не физическое явление. И не вычисление, и не модель, чем бы они ни были. Это нефизическое свойство внутри меня, благодаря которому яблоко делается ярким, присутствующим для меня, и я могу на него реагировать”.

Машина не только утверждает, что у нее есть сознание, — она, как и большинство людей, отвергает механистическое его объяснение и верит в метафизическую теорию. Она так устроена, она в ловушке своих собственных внутренних данных. Полагаясь на интроспекцию (под этим я понимаю именно когнитивную поисковую систему, обращающуюся к внутренним моделям машины), она неизбежно будет приходить к тому же метафизическому самоописанию.

О зрительном восприятии машин известно, пожалуй, достаточно, чтобы построить примерно такую несложную систему с использованием сегодняшних технологий. Пока они не продвинутся вперед, выйти за пределы зрения будет сложно, но в принципе ту же самую логику уместно применить к любому явлению, с которым столкнется робот. Теоретически можно заменить яблоко звуком, прикосновением, воспоминанием, эмоцией или мыслью, что дважды два четыре. Можно схематически выстроить мысль: “Я осознаю себя”, или даже: “Я осознаю, что сознаю”. Осознание яблока — только начало.

Примечания

Глава 1. Слон в комнате

1. Мои исследования личного пространства и сложных движений изложены в двух книгах. M. S. A. Graziano, *The Intelligent Movement Machine* (Oxford, UK: Oxford University Press, 2008); M. S. A. Graziano, *The Spaces between Us: A Story of Neuroscience, Evolution, and Human Nature* (New York: Oxford University Press, 2018).

2. Следующие ссылки дают комплексное представление об этой теории. Другие, более технические или сосредоточенные на экспериментальных данных работы здесь не приводятся. M. S. A. Graziano and S. Kastner, “Human Consciousness and Its Relationship to Social Neuroscience: A Novel Hypothesis,” *Cognitive Neuroscience* 2 (2011): 98–113; M. S. A. Graziano, *Consciousness and the Social Brain* (Oxford, UK: Oxford University Press, 2013); T. W. Webb and M. S. A. Graziano, “The Attention Schema Theory: A Mechanistic Account of Subjective Awareness,” *Frontiers in Psychology* 6 (2015): article 500.

3. Невозможно воздать здесь должное всем новым работам, посвященным механистическому, недуалистическому подходу к сознанию. Я привожу лишь несколько примеров и прошу прощения у множества блестящих авторов, которых не упоминаю. S. J. Blackmore, “Consciousness In Meme Machines,” *Journal of Consciousness Studies* 10 (2003): 19–30; P. S. Churchland, *Touching a Nerve: Our Brains, Our Selves* (New York: W. W. Norton, 2013); F. Crick, *The Astonishing Hypothesis: The Scientific Search for the Soul* (New York: Scribner, 1995); S. Dehaene, *Consciousness and the Brain* (New York: Viking Press, 2014); D. Dennett, *Consciousness Explained* (Boston: Back Bay Books, 1991); K. Frankish, “Illusionism as a Theory of Consciousness,” *Journal of Consciousness Studies* 23 (2016): 11–39; R. J. Gennaro, *Consciousness and Self Consciousness: A Defense of the Higher Order Thought Theory of Consciousness* (Philadelphia: John Benjamin’s Publishing, 1996); O. Holland and R. Goodman, “Robots with Internal Models: A Route to Machine Consciousness?” *Journal of Consciousness Studies* 10 (2003): 77–109; T. Metzinger, *The Ego Tunnel: The Science of the Mind and the Myth of the Self* (New York: Basic Books, 2009).

4. D. Chalmers, “Facing Up to the Problem of Consciousness,” *Journal of Consciousness Studies* 2 (1995): 200–219.

5. Более ранний и весьма прозорливый подход к сознанию, делающий акцент на внутренних моделях, содержится в работе: O. Holland and R. Goodman, “Robots with Internal Models: A Route

to Machine Consciousness?” *Journal of Consciousness Studies* 10 (2003): 77–109.

6. G. Ryle, *The Concept of Mind* (Chicago: University of Chicago Press, 1949).

7. J. Joyce, *Ulysses* (Paris: Sylvia Beach, 1922). Русский перевод: Джойс Д. Улисс / Пер. с англ. В. Хинкиса и С. Хоружего; коммент. С. Хоружего. — М.: Республика, 1993.

8. D. Chalmers, *The Character of Consciousness* (New York: Oxford University Press, 2010); T. Nagel, “What Is It Like to Be a Bat?” *The Philosophical Review* 83 (1974): 435–50; J. R. Searle, “Consciousness,” *Annual Review of Neuroscience* 23 (2000): 557–78.

9. R. A. Koene, “Scope and Resolution in Neural Prosthetics and Special Concerns for the Emulation of a Whole Brain,” *Journal of Geoethical Nanotechnology* 1 (2006): 21–29; R. Kurzweil, *The Singularity Is Near: When Humans Transcend Biology* (New York: Penguin Books, 2006); H. Markram, E. Muller, S. Ramaswamy, M. W. Reimann, M. Abdellah, C. A. Sanchez, A. Ailamaki, et al., “Reconstruction and Simulation of Neocortical Microcircuitry,” *Cell* 163 (2015): 456–92; A. Sandberg and N. Bostrom, “Whole Brain Emulation: A Roadmap,” Technical Report #2008–3, Future of Humanity Institute, Oxford University, 2008.

Глава 2. Крабы и осьминоги

1. И другие авторы убедительно описывали возможный ход эволюции сознания, включая туда связи сознания с вниманием (хотя делали это иначе, чем я). К примеру: C. Montemayor and H. H. Haladjian, *Consciousness, Attention, and Conscious Attention* (Cambridge, MA: MIT Press, 2015); R. Ornstein, *Evolution of Consciousness: The Origins of the Way We Think* (New York: Simon & Schuster, 1991).

2. O. Sakarya, K. A. Armstrong, M. Adamska, M. Adamski, I. F. Wang, B. Tidor, B. M. Degnan, T. H. Oakley, and K. S. Kosik, “A Post-Synaptic Scaffold at the Origin of the Animal Kingdom,” *PLoS One* 2 (2007): e506.

3. Z. Yin, M. Zhu, E. H. Davidson, D. J. Bottjer, F. Zhao, and P. Tafforeau, “Sponge Grade Body Fossil with Cellular Resolution Dating 60 Myr before the Cambrian,” *Proceedings of the National Academy of Sciences USA* 112 (2015): E1453–60.

4. D. H. Erwin, M. Laflamme, S. M. Tweedt, E. A. Sperling, D. Pisani, and K. J. Peterson, “The Cambrian Conundrum: Early Divergence and Later Ecological Success in the Early History of Animals,” *Science* 334 (2011): 1091–7; A. C. Marques and A. G. Collins, “Cladistic Analysis of Medusozoa and Cnidarian Evolution,” *Invertebrate Biology* 123 (2004): 23–42.

5. H. R. Bode, S. Heimfeld, O. Koizumi, C. L. Littlefield, and M. S. Yaross, “Maintenance and Regeneration of the Nerve Net in Hydra,” *American Zoology* 28 (1988): 1053–63.

6. R. B. Barlow Jr. and A. J. Fraioli, "Inhibition in the Limulus Lateral Eye in Situ," *Journal of General Physiology* 71 (1978): 699–720.

7. K. Hadeler, "On the Theory of Lateral Inhibition," *Kybernetik* 14 (1974): 161–5.

8. S. Koenemann and R. Jenner, *Crustacea and Arthropod Relationships* (Boca Raton: CRC Press, 2005).

9. B. Schoenemann, H. Pärnaste, and E. N. K. Clarkson, "Structure and Function of a Compound Eye, More Than Half a Billion Years Old," *Proceedings of the National Academy of Sciences USA* 114 (2017): 13489–94.

10. R. Gillette and J. W. Brown, "The Sea Slug, *Pleurobranchaea californica*: A Signpost Species in the Evolution of Complex Nervous Systems and Behavior," *Integrative and Comparative Biology* 55 (2015): 1058–69.

11. C. R. Smarandache-Wellmann, "Arthropod Neurons and Nervous System," *Current Biology* 26 (2016): R960–R965.

12. S. Koenig, R. Wolf, and M. Heisenberg, "Visual Attention in Flies — Dopamine in the Mushroom Bodies Mediates the After-Effect of Cueing," *PLoS One* 11 (2016): e0161412; B. van Swinderen, "Attention in *Drosophila*," *International Review of Neurobiology* 99 (2011): 51–85.

13. D. H. Erwin, M. Laflamme, S. M. Tweedt, E. A. Sperling, D. Pisani, and K. J. Peterson, "The Cambrian Conundrum: Early Divergence and Later Ecological Success in the Early History of Animals," *Science* 334 (211): 1091–97; B. Runnegar and J. Pojeta Jr., "Molluscan Phylogeny: The Paleontological Viewpoint," *Science* 186 (1974): 311–17.

14. J. Kluessendorf and P. Doyle, "*Pohlsepia mazonensis*, an Early 'Octopus' from the Carboniferous of Illinois, USA," *Palaeontology* 43 (2000): 919–26; A. R. Tanner, D. Fuchs, I. E. Winkelmann, M. T. Gilbert, M. S. Pankey, A. M. Ribeiro, K. M. Kocot, K. M. Halanych, T. H. Oakley, R. R. da Fonseca, D. Pisani, and J. Vinther, "Molecular Clocks Indicate Turnover and Diversification of Modern Coleoid Cephalopods during the Mesozoic Marine Revolution," *Proceedings of Royal Society, B, Biological Sciences* 284 (2017): 20162818.

15. P. Godfrey-Smith, *Other Minds: The Octopus, the Sea, and the Deep Origins of Consciousness* (New York: Farrar, Straus and Giroux, 2016); S. Montgomery, *The Soul of an Octopus* (New York: Atria Books, 2015).

16. A.-S. Darmaillacq, L. Dickel, and J. A. Mather, *Cephalopod Cognition* (Cambridge, UK: Cambridge University Press, 2014); D. B. Edelman, B. J. Baars, and A. K. Seth, "Identifying Hallmarks of Consciousness in Non-Mammalian Species," *Consciousness and Cognition* 14 (2015): 169–87; J. N. Richter, B. Hochner, and M. J. Kuba, "Pull or Push? Octopuses Solve a Puzzle Problem," *PLoS One* 11 (2016): e0152048.

17. B. Hochner, "An Embodied View of Octopus Neurobiology," *Current Biology* 22 (2012): R887–92.

18. P. M. Merikle, D. Smilek, and J. D. Eastwood, “Perception without Awareness: Perspectives from Cognitive Psychology,” *Cognition* 79 (2001): 115–34; R. Szczepanowski and L. Pessoa, “Fear Perception: Can Objective and Subjective Awareness Measures Be Dissociated?” *Journal of Vision* 10 (2007): 1–17.

Глава 3. Централизованный интеллект лягушки

1. E. Knudsen and J. S. Schwartz, “The Optic Tectum, a Structure Evolved for Stimulus Selection,” in *Evolution of Nervous Systems*, ed. J. Kaas (San Diego: Academic Press, 2017), 387–408; C. Maximino, “Evolutionary Changes in the Complexity of the Tectum of Nontetrapods: A Cladistic Approach,” *PLoS One* 3 (2008): e3582.
2. D. Ingle, “Visuomotor Functions of the Frog Optic Tectum,” *Brain, Behavior, and Evolution* 3 (1970): 57–71.
3. R. W. Sperry, “Effect of 180 Degree Rotation of the Retinal Field on Visuomotor Coordination,” *Journal of Experimental Zoology Part A: Ecological and Integrative Physiology* 92 (1943): 263–79; R. W. Sperry, “Optic nerve regeneration with return of vision in anurans,” *Journal of neurophysiology* 7.1 (1944): 57–69 (дополнение науч. ред.).
4. C. Comer and P. Grobstein, “Organization of Sensory Inputs to the Midbrain of the Frog, *Rana pipiens*,” *Journal of Comparative Physiology* 142 (1981): 161–68.
5. B. E. Stein and M. A. Meredith, *The Merging of the Senses* (Cambridge, MA: MIT Press, 1993).
6. C. Comer and P. Grobstein, “Organization of Sensory Inputs to the Midbrain of the Frog, *Rana pipiens*,” *Journal of Comparative Physiology* 142 (1981): 161–68; D. Ingle, “Visuomotor Functions of the Frog Optic Tectum,” *Brain, Behavior, and Evolution* 3 (1970): 57–71.
7. B. E. Stein and M. A. Meredith, *The Merging of the Senses* (Cambridge, MA: MIT Press, 1993).
8. T. Finkenstadt and J.-P. Ewert, “Visual Pattern Discrimination through Interactions of Neural Networks: A Combined Electrical Brain Stimulation, Brain Lesion, and Extracellular Recording Study in *Salamandra salamandra*,” *Journal of Comparative Physiology* 153 (1983): 99–110.
9. B. E. Stein and N. S. Gaither, “Sensory Representation in Reptilian Optic Tectum: Some Comparisons with Mammals,” *Journal of Comparative Neurology* 202 (1981): 69–87.
10. H. Vanegas and H. Ito, “Morphological Aspects of the Teleostean Visual System: A Review,” *Brain Research* 287 (1983): 117–37.
11. P. H. Hartline, L. Kass, and M. S. Loop, “Merging of Modalities in the Optic Tectum: Infrared and Visual Integration in Rattlesnakes,” *Science* 199 (1978): 1225–29.
12. S. P. Mysore and E. I. Knudsen, “The Role of a Midbrain Network in Competitive Stimulus Selection,” *Current Opinion in Neurobiology* 21 (2011): 653–60.

13. R. H. Wurtz and J. E. Albano, “Visual-Motor Function of the Primate Superior Colliculus,” *Annual Review of Neuroscience* 3 (1980): 189–226.

14. M. I. Posner, “Orienting of Attention,” *Quarterly Journal of Experimental Psychology* 32 (1980): 3–25.

15. E. F. Camacho and C. Bordons Alba, *Model Predictive Control* (New York: Springer, 2004); R. C. Conant and W. R. Ashby, “Every Good Regulator of a System Must Be a Model of That System,” *International Journal of Systems Science* 1 (1970): 89–97; B. A. Francis and W. M. Wonham, “The Internal Model Principle of Control Theory,” *Automatica* 12 (1976): 457–65.

16. M. S. A. Graziano and M. M. Botvinick, “How the Brain Represents the Body: Insights from Neurophysiology and Psychology,” in *Common Mechanisms in Perception and Action: Attention and Performance XIX*, ed. W. Prinz and B. Hommel (Oxford, UK: Oxford University Press, 2002), 136–57; N. Holmes and C. Spence, “The Body Schema and the Multisensory Representation (s) of Personal Space,” *Cognitive Processing* 5 (2004): 94–105; F. de Vignemont, *Mind the Body: An Exploration of Bodily Self-Awareness* (Oxford, UK: Oxford University Press, 2018).

17. H. Head and G. Holmes, “Sensory Disturbances from Cerebral Lesions,” *Brain* 34 (1911): 102–254; G. Vallar and R. Ronchi, “Somatoparaphrenia: A Body Delusion. A Review of the Neuropsychological Literature,” *Experimental Brain Research* 192 (2009): 533–51.

18. A. M. Haith and J. W. Krakauer, “Model-Based and Model-Free Mechanisms of Human Motor Learning,” in *Progress in Motor Control: Neural Computational and Dynamic Approaches, Volume 782*, ed. M. Richardson, M. Riley, and K. Shockley (New York: Springer, 2013), 1–21; S. M. McDougle, K. M. Bond, and J. A. Taylor, “Explicit and Implicit Processes Constitute the Fast and Slow Processes of Sensorimotor Learning,” *Journal of Neuroscience* 35 (2015): 9568–79; R. Shadmehr and F. A. Mussa-Ivaldi, “Adaptive Representation of Dynamics during Learning of a Motor Task,” *Journal of Neuroscience* 14 (1994): 3208–24.

19. Существует огромное количество экспериментальных работ, в которых изучался верхний холмик у кошек и обезьян, — в том числе то, как он отслеживает и прогнозирует положение головы и глаз, а следовательно, предсказывает, как зрительные образы будут двигаться по сетчатке. Я приведу лишь несколько обзорных статей. M. A. Basso and P. J. May, “Circuits for Action and Cognition: A View from the Superior Colliculus,” *Annual Review of Vision Science* 3 (2017): 197–226; D. L. Sparks, “Conceptual Issues Related to the Role of the Superior Colliculus in the Control of Gaze,” *Current Opinion in Neurobiology* 9 (1999): 698–707; R. H. Wurtz and J. E. Albano, “Visual-Motor Function of the Primate Superior Colliculus,” *Annual Review of Neuroscience* 3 (1980): 189–226.

Глава 4. Кора головного мозга и сознание

1. L. Medina and A. Reiner, “Do Birds Possess Homologues of Mammalian Primary Visual, Somatosensory and Motor Cortices?” *Trends in Neurosciences* 23 (2000): 1–12; R. K. Naumann and G. Laurent, “Function and Evolution of the Reptilian Cerebral Cortex,” in *Evolution of Nervous Systems*, ed. J. Kaas (San Diego: Academic Press, 2017), 491–518.

2. R. R. Lemon, *Vanished Worlds: An Introduction to Historical Geology* (Dubuque, IA: William C. Brown, 1993).

3. J. F. Harrison, A. Kaiser, and J. M. VandenBrooks, “Atmospheric Oxygen Level and the Evolution of Insect Body Size,” *Proceedings: Biological Science* 277 (2010): 1937–46.

4. R. L. Carroll, “The Origin and Early Radiation of Terrestrial Vertebrates,” *Journal of Paleontology* 75 (2001): 1202–13.

5. S. Sahney, M. J. Benton, and H. J. Falcon-Lang, “Rainforest Collapse Triggered Pensylvanian Tetrapod Diversification in Euramerica,” *Geology* 38 (2010): 1079–82.

6. R. K. Naumann and G. Laurent, “Function and Evolution of the Reptilian Cerebral Cortex,” in *Evolution of Nervous Systems*, ed. J. Kaas (San Diego: Academic Press, 2017), 491–518.

7. M. Leal and B. J. Powell, “Behavioural Flexibility and Problem-Solving in a Tropical Lizard,” *Biological Letters* 8 (2012): 28–30; J. D. Manrod, R. Hartdegen, and G. M. Burghardt, “Rapid Solving of a Problem Apparatus by Juvenile Black-Throated Monitor Lizards (*Varanus albigularis albigularis*),” *Animal Cognition* 11 (2008): 267–73; R. T. Mason and M. R. Parker, “Social Behavior and Pheromonal Communication in Reptiles,” *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology* 196 (2010): 729–49.

8. T. S. Kemp, *The Origin and Evolution of Mammals* (Oxford, UK: Oxford University Press, 2005); A. S. Romer and L. W. Price, “Review of the Pelycosauria,” *Geological Society of America, Special Papers* 28 (1940): 1–534.

9. Z. Molnár, J. H. Kaas, J. A. de Carlos, R. F. Hevner, E. Lein, and P. Němec, “Evolution and Development of the Mammalian Cerebral Cortex,” *Brain, Behavior, and Evolution* 83 (2014): 126–39.

10. A. B. Butler, “Evolution of the Thalamus: A Morphological and Functional Review,” *Thalamus and Related Systems* 4 (2008): 35–58; E. G. Jones, *The Thalamus* (New York: Springer, 1985).

11. P. Senter, “Phylogenetic Taxonomy and the Names of the Major Archosaurian (Reptilia) Clades,” *PaleoBios* 25 (2005): 1–7.

12. V. Dinets, “Apparent Coordination and Collaboration in Cooperatively Hunting Crocodilians,” *Ethology, Ecology, and Evolution* 27 (2012): 244–50; J. S. Doody, G. M. Burghardt, and V. Dinets, “Breaking the Social-Non-Social Dichotomy: A Role for Reptiles in Vertebrate Social Behavior Research?” *Ethology* 119 (2012): 1–9; L. D. Garrick and J. W.

Lang, "Social Signals and Behaviors of Adult Alligators and Crocodiles," *American Zoologist* 17 (1977): 225–39.

13. M. C. Langer, M. D. Ezcurra, J. S. Bittencourt, and F. E. Novas, "The Origin and Early Evolution of Dinosaurs," *Biological Reviews* 85 (2010): 55–110.

14. M. Bronzati, O. W. M. Rauhut, J. S. Bittencourt, and M. C. Langer, "Endocast of the Late Triassic (Carnian) Dinosaur *Saturnalia tupiniquim*: Implications for the Evolution of Brain Tissue in Sauropodomorpha," *Scientific Reports* 7 (2017): 11931; S. W. Rogers, "Allosaurus, Crocodiles, and Birds: Evolutionary Clues from Spiral Computed Tomography of an Endocast," *Anatomical Record* 257 (1999): 162–73.

15. L. M. Witmer and R. C. Ridgely, "New Insights into the Brain, Braincase, and Ear Region of Tyrannosaurs (Dinosauria, Theropoda), with Implications for Sensory Organization and Behavior," *Anatomical Record* 292 (2009): 1266–96.

16. S. L. Brusatte, J. K. O'Connor, and E. D. Jarvis, "The Origin and Diversification of Birds," *Current Biology* 25 (2015): R888–R898; L. M. Chiappe, *Glorified Dinosaurs: The Origin and Early Evolution of Birds* (Hoboken, NJ: John Wiley & Sons, 2007); Z. Zhou, "The Origin and Early Evolution of Birds: Discoveries, Disputes, and Perspectives from Fossil Evidence," *Naturwissenschaften* 91 (2004): 455–71.

17. Q. Ji and S. Ji, "On the Discovery of the Earliest Bird Fossil in China (*Sinosauropteryx*) and the Origin of Birds," *Chinese Geology* 10 (1996): 30–33; M. A. Norell and X. Xu, "Feathered Dinosaurs," *Annual Review of Earth and Planetary Science* 33 (2005): 277–99.

18. H. J. Karten, "Vertebrate Brains and Evolutionary Connectomics: On the Origins of the Mammalian 'Neocortex,'" *Philosophical Transactions of the Royal Society of London, B, Biological Sciences* 370 (2015): 20150060; L. Medina and A. Reiner, "Do Birds Possess Homologues of Mammalian Primary Visual, Somatosensory and Motor Cortices?" *Trends in Neurosciences* 23 (2000): 1–12.

19. G. R. Hunt and R. D. Gray, "Tool Manufacture by New Caledonian Crows: Chipping Away at Human Uniqueness," *Acta Zoologica Sinica (Supplement)* 52 (2006): 622–25; C. Rutz and J. J. St Clair, "The Evolutionary Origins and Ecological Context of Tool Use in New Caledonian Crows," *Behavioral Processes* 89 (2012): 153–65; C. Rutz, S. Sugasawa, J. E. van der Wal, B. C. Klump, and J. J. St Clair, "Tool Bending in New Caledonian Crows," *Royal Society of Open Science* 3 (2016): 160439.

20. S. A. Jelbert, A. H. Taylor, L. G. Cheke, N. S. Clayton, and R. D. Gray, "Using the Aesop's Fable Paradigm to Investigate Causal Understanding of Water Displacement by New Caledonian Crows," *PLoS One* 9 (2014): e92895.

21. D. M. Beck and S. Kastner, "Top-Down and Bottom-Up Mechanisms in Biasing Competition in the Human Brain," *Vision Research* 49 (2009): 1154–65; R. Desimone and J. Duncan, "Neural Mechanisms of

Selective Visual Attention," *Annual Review of Neuroscience* 18 (1995): 193–222.

22. Существует огромное количество литературы, посвященной мозаике зрительных зон в коре головного мозга приматов. Над этим работали тысячи людей, в том числе и я. Приведу лишь несколько полезных источников, описывающих исследования как обезьян, так и людей. D. Felleman and D. Van Essen, "Distributed Hierarchical Processing in the Primate Visual Cortex," *Cerebral Cortex* 1 (1991): 1–47; K. Grill-Spector and R. Malach, "The Human Visual Cortex," *Annual Review of Neuroscience* 27 (2004): 649–77; P. Schiller and E. Tehovnik, *Vision and the Visual System* (Oxford, UK: Oxford University Press, 2015); L. G. Ungerleider and J. V. Haxby, "'What' and 'Where' in the Human Brain," *Current Opinion in Neurobiology* 4 (1994): 157–65; D. C. Van Essen, J. W. Lewis, H. A. Drury, N. Hadjikhani, R. B. Tootell, M. Bakircioglu, and M. I. Miller, "Mapping Visual Cortex in Monkeys and Humans Using Surface-Based Atlases," *Vision Research* 41 (2001): 1359–78; L. Wang, R. E. Mruczek, M. J. Arcaro, and S. Kastner, "Probabilistic Maps of Visual Topography in Human Cortex," *Cerebral Cortex* 25 (2015): 3911–31.

23. T. Moore and M. Zirnsak, "Neural Mechanisms of Selective Visual Attention," *Annual Review of Psychology* 68 (2017): 47–72.

24. R. Desimone and J. Duncan, "Neural Mechanisms of Selective Visual Attention," *Annual Review of Neuroscience* 18 (1995): 193–222.

25. G. Alarcon and A. Valentin, *Introduction to Epilepsy* (Cambridge, UK: Cambridge University Press, 2012).

26. R. B. Barlow Jr. and A. J. Fraioli, "Inhibition in the Limulus Lateral Eye in Situ," *Journal of General Physiology* 71 (1978): 699–720; K. Hadeler, "On the Theory of Lateral Inhibition," *Kybernetik* 14 (1974): 161–65.

27. M. Corbetta, G. Patel, and G. L. Shulman, "The Reorienting System of the Human Brain: From Environment to Theory of Mind," *Neuron* 58 (2008): 306–24; K. Igelström and M. S. A. Graziano, "The Inferior Parietal Lobe and Temporoparietal Junction: A Network Perspective," *Neuropsychologia* 105 (2017): 70–83; R. Saxe and L. J. Powell, "It's the Thought That Counts: Specific Brain Regions for One Component of Theory of Mind," *Psychological Science* 17 (2006): 692–9; M. Scolari, K. N. Seidl-Rathkopf, and S. Kastner, "Functions of the Human Frontoparietal Attention Network: Evidence from Neuroimaging," *Current Opinion in Behavioral Sciences* 1 (2015): 32–39; J. L. Vincent, I. Kahn, A. Z. Snyder, M. E. Raichle, and R. L. Buckner, "Evidence for a Frontoparietal Control System Revealed by Intrinsic Functional Connectivity," *Journal of Neurophysiology* 100 (2008): 3328–42; B. T. T. Yeo, F. M. Krienen, J. Sepulcre, M. R. Sabuncu, D. Lashkari, M. Hollinshead, J. L. Roffman, et al., "The Organization of the Human Cerebral Cortex Estimated by Intrinsic Functional Connectivity," *Journal of Neurophysiology* 106 (2011): 1125–65.

28. D. Dennett, *Sweet Dreams: Philosophical Obstacles to a Science of Consciousness (Jean Nicod Lectures)* (Cambridge, MA: MIT Press,

2005). Русский перевод: Д. Деннет. Сладкие грёзы: Чем философия мешает науке о сознании / Под ред. М. О. Кедровой; пер. с англ. А. Н. Коваля. — М.: URSS, 2017.

29. C. L. Colby and M. E. Goldberg, “Space and Attention in Parietal Cortex,” *Annual Review of Neuroscience* 22 (1999): 319–49; J. Gottlieb, “From Thought to Action: The Parietal Cortex as a Bridge between Perception, Action, and Cognition,” *Neuron* 53 (2007): 9–16; E. J. Tehovnik, M. A. Sommer, I. H. Chou, W. M. Slocum, and P. H. Schiller, “Eye Fields in the Frontal Lobes of Primates,” *Brain Research Reviews* 32 (2000): 413–48.

30. C. Eriksen and J. St James, “Visual Attention within and around the Field of Focal Attention: A Zoom Lens Model,” *Perception and Psychophysics* 40 (1986): 225–40; M. I. Posner, C. R. Snyder, and B. J. Davidson, “Attention and the Detection of Signals,” *Journal of Experimental Psychology* 109 (1980): 160–74.

31. M. Scolari, E. F. Ester, and J. T. Serences, “Feature- and Object-Based Attentional Modulation in the Human Visual System,” in *The Oxford Handbook of Attention*, ed. A. C. Norbre and S. Kastner (Oxford, UK: Oxford University Press, 2015), 573–600; S. Treue, “Objectand Feature-Based Attention: Monkey Physiology,” in *The Oxford Handbook of Attention*, ed. A. C. Norbre and S. Kastner (Oxford, UK: Oxford University Press, 2015), 601–19.

32. Ученые так часто говорили о связи сложности и сознания, что эта мысль прижилась в научной фантастике. Джулио Тонони изложил эту гипотезу в самой системной, математической форме. G. Tononi, *Phi: A Voyage from the Brain to the Soul* (New York: Pantheon, 2012).

33. R. Bshary, W. Wickler, and H. Fricke, “Fish Cognition: A Primate’s Eye View,” *Animal Cognition* 5 (2002): 1–13.

34. C. Koch, “Consciousness Redux: What Is It Like to Be a Bee?” *Scientific American Mind* 19 (December 2008): 18–19.

35. D. Skrbina, *Panpsychism in the West* (Boston: MIT Press, 2005).

36. B. J. Baars, *A Cognitive Theory of Consciousness* (New York: Cambridge University Press, 1988); S. Dehaene, *Consciousness and the Brain* (New York: Viking Press, 2014).

37. E. Todorov and M. I. Jordan, “Optimal Feedback Control as a Theory of Motor Coordination,” *Nature Neuroscience* 5 (2002): 1226–35.

Глава 5. Социальное сознание

1. M. J. Doherty, *How Children Understand Others’ Thoughts and Feelings* (New York: Psychology Press, 2008); U. Frith and C. D. Frith, “Development and Neurophysiology of Mentalizing,” *Philosophical Transactions of the Royal Society of London, B, Biological Sciences* 358 (2003): 459–73; D. Premack and G. Woodruff, “Does the Chimpanzee Have a Theory of Mind?” *Behavioral and Brain Sciences* 1 (1978): 515–26.

2. S. Baron-Cohen, *Mindblindness: An Essay on Autism and Theory of Mind* (Cambridge, MA: MIT Press, 1997); C. K. Friesen and A. Kingstone, “The Eyes Have it! Reflexive Orienting Is Triggered by Nonpredictive Gaze,” *Psychonomic Bulletin and Review* 5 (1998): 490–95; E. A. Hoffman and J. V. Haxby, “Distinct Representations of Eye Gaze and Identity in the Distributed Human Neural System for Face Perception,” *Nature Neuroscience* 3 (2000): 80–84; L. A. Symons, K. Lee, C. C. Cedrone, and M. Nishimura, “What Are You Looking At? Acuity for Triadic Eye Gaze,” *Journal of General Psychology* 131 (2004): 451–69.

3. S. Baron-Cohen, A. M. Leslie, and U. Frith, “Does the Autistic Child Have a ‘Theory of Mind’?” *Cognition* 21 (1985): 37–46; H. Wimmer and J. Perner, “Beliefs about Beliefs: Representation and Constraining Function of Wrong Beliefs in Young Children’s Understanding of Deception,” *Cognition* 13 (1983): 103–28.

4. H. M. Wellman, D. Cross, and J. Watson, “Meta-Analysis of Theory-of-Mind Development: The Truth about False Belief,” *Child Development* 72 (2001): 655–84.

5. C. Krupenye, F. Kano, S. Hirata, J. Call, and M. Tomasello, “Great Apes Anticipate That Other Individuals Will Act According to False Beliefs,” *Science* 354 (2016): 110–14.

6. N. S. Clayton, “Ways of Thinking: From Crows to Children and Back Again,” *Quarterly Journal of Experimental Psychology* 68 (2015): 209–41.

7. J. J. Gibson, *The Ecological Approach to Visual Perception* (Boston: Houghton Mifflin Harcourt, 1979).

8. D. C. Dennett, *The Intentional Stance* (Cambridge, MA: Bradford Books/MIT Press, 1987).

9. C. L. Baker, R. Saxe, and J. B. Tenenbaum, “Action Understanding as Inverse Planning,” *Cognition* 113 (2009): 329–49; N. C. Rabinowitz, F. Perbet, F. Song, C. Zhang, S. M. Ali Eslami, and M. Botvinick, “Machine Theory of Mind,” *Computer Science arXiv* (2017): 1802.007740; R. Saxe and S. D. Houlihan, “Formalizing Emotion Concepts within a Bayesian Model of Theory of Mind,” *Current Opinion in Psychology* 17 (2017): 15–21.

10. D. J. Acheson, *Elementary Fluid Dynamics* (Oxford, UK: Clarendon Press, 1990).

11. A. Guterstam, H. H. Kean, T. W. Webb, F. S. Kean, and M. S. A. Graziano, “An Implicit Model of Other People’s Visual Attention as an Invisible, Force-Carrying Beam Projecting from the Eyes,” *Proceedings of the National Academy of Sciences USA* (in press).

12. C. G. Gross, “The Fire That Comes from the Eye,” *The Neuroscientist* 5 (1999): 58–64.

13. A. Dundes, *The Evil Eye: A Folklore Casebook* (New York: Garland Press, 1981).

14. E. B. Titchener, “The Feeling of Being Stared At,” *Science* 8 (1898): 895–7.

15. J. Piaget, *The Child’s Conception of the World*, trans. J. Tomlinson and A. Tomlinson (Totowa, NJ: Little, Adams, 1979).

16. G. A. Winer, J. E. Cottrell, V. Gregg, J. S. Fournier, and L. S. Bica, “Fundamentally Misunderstanding Visual Perception: Adults’ Belief in Visual Emissions,” *American Psychologist* 57 (2002): 417–24; G. A. Winer, J. E. Cottrell, and K. D. Karefilaki, “Images, Words and Questions: Variables That Influence Beliefs about Vision in Children and Adults,” *Journal of Experimental Child Psychology* 63 (1996): 499–525.

Глава 6. Мастер Йода и Дарт Вейдер: как найти в мозге сознание?

1. J. H. Kaas, “The Evolution of Brains from Early Mammals to Humans,” *Wiley Interdisciplinary Review of Cognitive Science* 4 (2013): 33–45.

2. J. E. Bogen, “Some Neurophysiologic Aspects of Consciousness,” *Seminars in Neurology* 17 (1997): 95–103; G. M. Edelman, J. A. Gally, and B. J. Baars, “Biology of Consciousness,” *Frontiers in Psychology* 2 (2011): 4; L. M. Ward, “The Thalamic Dynamic Core Theory of Conscious Experience,” *Consciousness and Cognition* 20 (2011): 464–86.

3. E. G. Jones, *The Thalamus* (New York: Springer, 1985).

4. F. C. Crick and C. Koch, “What Is the Function of the Claustrum?” *Philosophical Transactions of the Royal Society of London, B, Biological Sciences* 360 (2005): 1271–79; Y. Goll, G. Atlan, and A. Citri, “Attention: The Claustrum,” *Trends in Neurosciences* 38 (2015): 486–95; Z. K. Mohamad, B. Fabrice, B. Abdelrahman, and P. Fabienne, “Electrical Stimulation of a Small Brain Area Reversibly Disrupts Consciousness,” *Epilepsy and Behavior* 37 (2014): 32–35.

5. R. Blake, J. Brascamp, and D. J. Heeger, “Can Binocular Rivalry Reveal Neural Correlates of Consciousness?” *Philosophical Transactions of the Royal Society of London, B, Biological Sciences* 369 (2014): 20130211; R. Blake and N. K. Logothetis, “Visual Competition,” *Nature Reviews Neuroscience* 3 (2002): 13–21; D. A. Leopold and N. K. Logothetis, “Activity Changes in Early Visual Cortex Reflect Monkeys’ Percepts during Binocular Rivalry,” *Nature* 379 (1996): 549–53; B. A. Metzger, K. E. Mathewson, E. Tapia, M. Fabiani, G. Gratton, and D. M. Beck, “Regulating the Access to Awareness: Brain Activity Related to Probe-Related and Spontaneous Reversals in Binocular Rivalry,” *Journal of Cognitive Neuroscience* 29 (2017): 1089–102; K. Sandberg, B. Bahrami, R. Kanai, G. R. Barnes, M. Overgaard, and G. Rees, “Early Visual Responses Predict Conscious Face Perception within and between Subjects during Binocular Rivalry,” *Journal of Cognitive Neuroscience* 25 (2013): 969–85; F. Tong, M. Meng, and R. Blake, “Neural Bases of Binocular Rivalry,” *Trends in Cognitive Sciences* 10 (2006): 502–11.

6. R. Blake and N. K. Logothetis, “Visual Competition,” *Nature Reviews Neuroscience* 3 (2002): 13–21; D. A. Leopold and N. K. Logothetis,

"Activity Changes in Early Visual Cortex Reflect Monkeys' Percepts during Binocular Rivalry," *Nature* 379 (1996): 549–53.

7. R. Blake, J. Brascamp, and D. J. Heeger, "Can Binocular Rivalry Reveal Neural Correlates of Consciousness?" *Philosophical Transactions of the Royal Society of London, B, Biological Sciences* 369 (2014): 20130211; B. A. Metzger, K. E. Mathewson, E. Tapia, M. Fabiani, G. Gratton, and D. M. Beck, "Regulating the Access to Awareness: Brain Activity Related to Probe-Related and Spontaneous Reversals in Binocular Rivalry," *Journal of Cognitive Neuroscience* 29 (2017): 1089–102; K. Sandberg, B. Bahrami, R. Kanai, G. R. Barnes, M. Overgaard, and G. Rees, "Early Visual Responses Predict Conscious Face Perception within and between Subjects during Binocular Rivalry," *Journal of Cognitive Neuroscience* 25 (2013): 969–85; F. Tong, M. Meng, and R. Blake, "Neural Bases of Binocular Rivalry," *Trends in Cognitive Sciences* 10 (2006): 502–11.

8. R. Blake and N. K. Logothetis, "Visual Competition," *Nature Reviews Neuroscience* 3 (2002): 13–21; F. Tong, M. Meng, and R. Blake, "Neural Bases of Binocular Rivalry," *Trends in Cognitive Sciences* 10 (2006): 502–11.

9. K. Wunderlich, K. A. Schneider, and S. Kastner, "Neural Correlates of Binocular Rivalry in the Human Lateral Geniculate Nucleus," *Nature Neuroscience* 8 (2005): 1595–602.

10. M. S. Beauchamp, J. V. Haxby, J. E. Jennings, and E. A. DeYoe "An fMRI Version of the Farnsworth-Munsell 100-Hue Test Reveals Multiple Color-Selective Areas in Human Ventral Occipitotemporal Cortex," *Cerebral Cortex* 9 (1999): 257–63; B. R. Conway, "Color Signals through Dorsal and Ventral Visual Pathways," *Visual Neuroscience* 31 (2014): 197–209.

11. R. Blake and N. K. Logothetis, "Visual Competition," *Nature Reviews Neuroscience* 3 (2002): 13–21; H. H. Li, J. Rankin, J. Rinzel, M. Carrasco, and D. J. Heeger, "Attention Model of Binocular Rivalry," *Proceedings of the National Academy of Sciences USA* 114 (2017): E6192 — E6201; F. Tong, M. Meng, and R. Blake, "Neural Bases of Binocular Rivalry," *Trends in Cognitive Sciences* 10 (2006): 502–11.

12. R. Blake and N. K. Logothetis, "Visual Competition," *Nature Reviews Neuroscience* 3 (2002): 13–21; H. H. Li, J. Rankin, J. Rinzel, M. Carrasco, and D. J. Heeger, "Attention Model of Binocular Rivalry," *Proceedings of the National Academy of Sciences USA* 114 (2017): E6192 — E6201; F. Tong, M. Meng, and R. Blake, "Neural Bases of Binocular Rivalry," *Trends in Cognitive Sciences* 10 (2006): 502–11.

13. M. S. Beauchamp, J. V. Haxby, J. E. Jennings, and E. A. DeYoe, "An fMRI Version of the Farnsworth-Munsell 100-Hue Test Reveals Multiple Color-Selective Areas in Human Ventral Occipitotemporal Cortex," *Cerebral Cortex* 9 (1999): 257–63.

14. S. E. Bouvier and S. A. Engel, "Behavioral Deficits and Cortical Damage Loci in Cerebral Achromatopsia," *Cerebral Cortex* 16 (2006): 183–91.

15. M. Binder, K. Gociewicz, B. Windey, M. Koculak, K. Finc, J. Nikadon, M. Derda, and A. Cleeremans, "The Levels of Perceptual Processing and the Neural Correlates of Increasing Subjective Visibility," *Consciousness and Cognition* 55 (2017): 106–25; D. Carmel, N. Lavie, and G. Rees, "Conscious Awareness of Flicker in Humans Involves Frontal and Parietal Cortex," *Current Biology* 16 (2006): 907–11; M. S. Christensen, T. Z. Ramsøy, T. E. Lund, K. H. Madsen, and J. B. Rowe, "An fMRI Study of the Neural Correlates of Graded Visual Perception," *Neuroimage* 31 (2006): 1711–25; S. Dehaene and J. P. Changeux, "Experimental and Theoretical Approaches to Conscious Processing," *Neuron* 70 (2011): 200–227; S. Dehaene, L. Naccache, L. Cohen, D. L. Bihan, J. F. Mangin, J. B. Poline, and D. Rivière, "Cerebral Mechanisms of Word Masking and Unconscious Repetition Priming," *Nature Neuroscience* 4 (2001): 752–58.

16. A. Schurger, I. Sarigiannidis, L. Naccache, J. D. Sitt, and S. Dehaene, "Cortical Activity Is More Stable When Sensory Stimuli Are Consciously Perceived," *Proceedings of the National Academy of Sciences USA* 112 (2015): E2083–92.

17. M. Binder, K. Gociewicz, B. Windey, M. Koculak, K. Finc, J. Nikadon, M. Derda, and A. Cleeremans, "The Levels of Perceptual Processing and the Neural Correlates of Increasing Subjective Visibility," *Consciousness and Cognition* 55 (2017): 106–25; D. Carmel, N. Lavie, and G. Rees, "Conscious Awareness of Flicker in Humans Involves Frontal and Parietal Cortex," *Current Biology* 16 (2006): 907–11; M. S. Christensen, T. Z. Ramsøy, T. E. Lund, K. H. Madsen, and J. B. Rowe, "An fMRI Study of the Neural Correlates of Graded Visual Perception," *Neuroimage* 31 (2006): 1711–25; S. Dehaene and J. P. Changeux, "Experimental and Theoretical Approaches to Conscious Processing," *Neuron* 70 (2011): 200–227; S. Dehaene, L. Naccache, L. Cohen, D. L. Bihan, J. F. Mangin, J. B. Poline, and D. Rivière, "Cerebral Mechanisms of Word Masking and Unconscious Repetition Priming," *Nature Neuroscience* 4 (2001): 752–58.

18. T. W. Webb, K. Igelström, A. Schurger, and M. S. A. Graziano, "Cortical Networks Involved in Visual Awareness Independently of Visual Attention," *Proceedings of the National Academy of Sciences USA* 113 (2016): 13923–28.

19. T. J. Buschman and E. K. Miller, "Goal-Direction and Top-Down Control," *Philosophical Transactions of the Royal Society of London, B, Biological Sciences* 369 (2014): 20130471; E. K. Miller and J. D. Cohen, "An Integrative Theory of Prefrontal Cortex Function," *Annual Review of Neurosciences* 24 (2001): 167–202.

20. A. Nieder and E. K. Miller, "Coding of Cognitive Magnitude: Compressed Scaling of Numerical Information in the Primate Prefrontal Cortex," *Neuron* 37 (2003): 149–57.

21. S. C. Rao, G. Rainer, and E. K. Miller, "Integration of What and Where in the Primate Prefrontal Cortex," *Science* 276 (1997): 821–24.

22. D. J. Freedman, M. Riesenhuber, T. Poggio, and E. K. Miller, "Categorical Representation of Visual Stimuli in the Primate

Prefrontal Cortex," *Science* 291 (2001): 312–16.

23. R. Levy and P. S. Goldman-Rakic, "Segregation of Working Memory Functions within the Dorsolateral Prefrontal Cortex," *Experimental Brain Research* 133 (2000): 23–32; E. K. Miller, "The 'Working' of Working Memory," *Dialogues in Clinical Neuroscience* 15 (2013): 411–18.

24. B. Odegaard, R. T. Knight, and H. Lau, "Should a Few Null Findings Falsify Prefrontal Theories of Conscious Perception?" *Journal of Neuroscience* 40 (2017): 9593–602.

25. В нескольких солидных обзорных работах подчеркивается, что утрата или снижение сознания обычно не считается симптомом повреждения префронтальной коры. J. Fuster, *The Prefrontal Cortex* (New York: Academic Press, 2015); A. Henri-Bhargava, D. T. Stuss, and M. Freedman, "Clinical Assessment of Prefrontal Lobe Functions," *Continuum, Behavioral Neurology and Psychiatry* 24 (2018): 704–26; T. Shallice and L. Cipolotti, "The Prefrontal Cortex and Neurological Impairments of Active Thought," *Annual Review of Psychology* 69 (2018): 157–80; S. M. Szczepanski and R. T. Knight, "Insights into Human Behavior from Lesions to the Prefrontal Cortex," *Neuron* 83 (2014): 1002–18.

26. Работы по теменно-лобным сетям сыплются, как из рога изобилия. Вот необходимый минимум источников, где речь идет об основных сетях, упомянутых в книге. D. Bzdok, R. Langner, L. Schilbach, O. Jakobs, C. Roski, S. Caspers, A. R. Laird, et al. "Characterization of the Temporo-Parietal Junction by Combining Data-Driven Parcellation, Complementary Connectivity Analyses, and Functional Decoding," *Neuroimage* 81 (2013): 381–92; M. Corbetta, G. Patel, and G. L. Shulman, "The Reorienting System of the Human Brain: From Environment to Theory of Mind," *Neuron* 58 (2008): 306–24; N. U. Dosenbach, D. A. Fair, F. M. Miezin, A. L. Cohen, K. K. Wenger, R. A. Dosenbach, M. D. Fox, A. Z. Snyder, et al., "Distinct Brain Networks for Adaptive and Stable Task Control in Humans," *Proceedings of the National Academy of Sciences USA* 104 (2007): 11073–78; M. D. Fox, M. Corbetta, A. Z. Snyder, J. L. Vincent, and M. E. Raichle, "Spontaneous Neuronal Activity Distinguishes Human Dorsal and Ventral Attention Systems," *Proceedings of the National Academy of Sciences USA* 103 (2006): 10046–51; K. Igelström and M. S. A. Graziano, "The Inferior Parietal Lobe and Temporoparietal Junction: A Network Perspective," *Neuropsychologia* 105 (2017): 70–83; K. M. Igelström, T. W. Webb, and M. S. A. Graziano, "Neural Processes in the Human Temporoparietal Cortex Separated by Localized Independent Component Analysis," *Journal of Neuroscience* 35 (2015): 9432–45; K. M. Igelström, T. W. Webb, and M. S. A. Graziano, "Topographical Organization of Attentional, Social and Memory Processes in the Human Temporoparietal Cortex," *eNeuro* 3 (2016): e0060; R. B. Mars, J. Sallet, U. Schüffelgen, S. Jbabdi, I. Toni, and M. F. Rushworth, "Connectivity-Based Subdivisions of the Human Right Temporoparietal Junction Area: Evidence for Different Areas Participating in Different Cortical Networks," *Cerebral Cortex* 22 (2012): 1894–903; R. Ptak, "The Frontoparietal Attention Network of

the Human Brain: Action, Saliency, and a Priority Map of the Environment," *Neuroscientist* 18 (2012): 502–15; R. Saxe and L. J. Powell, "It's the Thought That Counts: Specific Brain Regions for One Component of Theory of Mind," *Psychological Science* 17 (2006): 692–99; M. Scolari, K. N. Seidl-Rathkopf, and S. Kastner, "Functions of the Human Frontoparietal Attention Network: Evidence from Neuroimaging," *Current Opinion in Behavioral Sciences* 1 (2015): 32–39; J. L. Vincent, I. Kahn, A. Z. Snyder, M. E. Raichle, and R. L. Buckner, "Evidence for a Frontoparietal Control System Revealed by Intrinsic Functional Connectivity," *Journal of Neurophysiology* 100 (2008): 3328–42; B. T. T. Yeo, F. M. Krienen, J. Sepulcre, M. R. Sabuncu, D. Lashkari, M. Hollinshead, J. L. Roffman, et al., "The Organization of the Human Cerebral Cortex Estimated by Intrinsic Functional Connectivity," *Journal of Neurophysiology* 106 (2011): 1125–65.

27. C. Amiez and M. Petrides, "Anatomical Organization of the Eye Fields in the Human and Non-Human Primate Frontal Cortex," *Progress in Neurobiology* 89 (2009): 220–30; L. L. Chen and E. J. Tehovnik, "Cortical Control of Eye and Head Movements: Integration of Movements and Percepts," *European Journal of Neuroscience* 25 (2007): 1253–64; M. H. Grosbras and A. Berthoz, "Parieto-Frontal Networks and Gaze Shifts in Humans: Review of Functional Magnetic Resonance Imaging Data," *Advances in Neurology* 93 (2003): 269–80; E. Lobel, P. Kahane, U. Leonards, M. Grosbras, S. Lehericy, D. Le Bihan, and A. Berthoz, "Localization of Human Frontal Eye Fields: Anatomical and Functional Findings of Functional Magnetic Resonance Imaging and Intracerebral Electrical Stimulation," *Journal of Neurosurgery* 95 (2001): 804–15.

28. R. Caminiti, S. Ferraina, and P. B. Johnson, "The Sources of Visual Information to the Primate Frontal Lobe: A Novel Role for the Superior Parietal Lobule," *Cerebral Cortex* 6 (1996): 319–28; C. S. Konen, R. E. Mruczek, J. L. Montoya, and S. Kastner, "Functional Organization of Human Posterior Parietal Cortex: Grasping- and Reaching-Related Activations Relative to Topographically Organized Cortex," *Journal of Neurophysiology* 109 (2013): 2897–908; L. H. Snyder, A. P. Batista, and R. A. Andersen, "Coding of Intention in the Posterior Parietal Cortex," *Nature* 386 (1997): 167–70.

29. M. G. Di Bono, C. Begliomini, U. Castiello, and M. Zorzi, "Probing the Reaching-Grasping Network in Humans through Multivoxel Pattern Decoding," *Brain and Behavior* 5 (2015): e00412; C. S. Konen, R. E. Mruczek, J. L. Montoya, and S. Kastner, "Functional Organization of Human Posterior Parietal Cortex: Gras-ping- and Reaching-Related Activations Relative to Topographically Organized Cor-tex," *Journal of Neurophysiology* 109 (2013): 2897–908; A. Murata, V. Gallese, G. Luppino, M. Kaseda, and H. Sakata, "Selectivity for the Shape, Size, and Orientation of Objects for Grasping in Neurons of Monkey Parietal Area AIP," *Journal of Neurophysiology* 83 (2000): 2580–601; G. Rizzolatti, R. Camarda, L. Fogassi, M. Gentilucci, G. Luppino, and M. Matelli, "Functional Organization of Inferior Area 6 in the Macaque

Monkey. II. Area F5 and the Control of Distal Movements," *Experimental Brain Research* 71 (1988): 491–507.

30. D. F. Cooke and M. S. A. Graziano, "Super-Flinchers and Nerves of Steel: Defensive Movements Altered by Chemical Manipulation of a Cortical Motor Area," *Neuron* 43 (2004): 585–93; D. F. Cooke, C. S. R. Taylor, T. Moore, and M. S. A. Graziano, "Complex Movements Evoked by Microstimulation of Area VIP," *Proceedings of the National Academy of Sciences USA* 100 (2003): 6163–68.

31. E. Eger, P. Pinel, S. Dehaene, and A. Kleinschmidt, "Spatially Invariant Coding of Numerical Information in Functionally Defined Subregions of Human Parietal Cortex," *Cerebral Cortex* 25 (2015): 1319–29; A. Nieder and E. K. Miller, "Coding of Cognitive Magnitude: Compressed Scaling of Numerical Information in the Primate Prefrontal Cortex," *Neuron* 37 (2003): 149–57; R. Stanescu-Cosson, P. Pinel, P. F. van De Moortele, D. Le Bihan, L. Cohen, and S. Dehaene, "Understanding Dissociations in Dyscalculia: A Brain Imaging Study of the Impact of Number Size on the Cerebral Networks for Exact and Approximate Calculation," *Brain* 123 (2000): 2240–55.

32. K. M. Igelström, T. W. Webb, and M. S. A. Graziano, "Topographical Organization of Attentional, Social and Memory Processes in the Human Temporoparietal Cortex," *eNeuro* 3 (2016): e0060; Y. T. Kelly, T. W. Webb, J. D. Meier, M. J. Arcaro, and M. S. A. Graziano, "Attributing Awareness to Oneself and to Others," *Proceedings of the National Academy of Sciences USA* 111 (2014): 5012–17; T. W. Webb, K. Igelström, A. Schurger, and M. S. A. Graziano, "Cortical Networks Involved in Visual Awareness Independently of Visual Attention," *Proceedings of the National Academy of Sciences USA* 113 (2016): 13923–28.

33. K. Igelström and M. S. A. Graziano, "The Inferior Parietal Lobe and Temporoparietal Junction: A Network Perspective," *Neuropsychologia* 105 (2017): 70–83; K. M. Igelström, T. W. Webb, and M. S. A. Graziano, "Topographical Organization of Attentional, Social and Memory Processes in the Human Temporoparietal Cortex," *eNeuro* 3 (2016): e0060; R. B. Mars, J. Sallet, U. Schüffelgen, S. Jbabdi, I. Toni, and M. F. S. Rushworth, "Connectivity-Based Subdivisions of the Human Right Temporoparietal Junction Area: Evidence for Different Areas Participating in Different Cortical Networks," *Cerebral Cortex* 22 (2012): 1894–903.

34. R. Saxe and L. J. Powell, "It's the Thought That Counts: Specific Brain Regions for One Component of Theory of Mind," *Psychological Science* 17 (2006): 692–99.

35. Y. T. Kelly, T. W. Webb, J. D. Meier, M. J. Arcaro, and M. S. A. Graziano, "Attributing Awareness to Oneself and to Others," *Proceedings of the National Academy of Sciences USA* 111 (2014): 5012–17; T. W. Webb, K. Igelström, A. Schurger, and M. S. A. Graziano, "Cortical Networks Involved in Visual Awareness Independently of Visual Attention," *Proceedings of the National Academy of Sciences USA* 113 (2016): 13923–28.

36. M. Corbetta, G. Patel, and G. L. Shulman, "The Reorienting System of the Human Brain: From Environment to Theory of Mind," *Neuron* 58 (2008): 306–24; T. Moore and M. Zirnsak, "Neural Mechanisms of Selective Visual Attention," *Annual Review of Psychology* 68 (2017): 47–72; R. Ptak, "The Frontoparietal Attention Network of the Human Brain: Action, Saliency, and a Priority Map of the Environment," *Neuroscientist* 18 (2012): 502–15.

37. K. M. Igelström, T. W. Webb, and M. S. A. Graziano, "Topographical Organization of Attentional, Social and Memory Processes in the Human Temporoparietal Cortex," *eNeuro* 3 (2016): e0060.

38. M. A. Goodale and A. D. Milner, "Separate Visual Pathways for Perception and Action," *Trends in Neurosciences* 15 (1992): 20–25.

39. M. Hurme, M. Koivisto, A. Revonsuo, and H. Railo, "Early Processing in Primary Visual Cortex Is Necessary for Conscious and Unconscious Vision While Late Processing Is Necessary Only for Conscious Vision in Neurologically Healthy Humans," *Neuroimage* 150 (2017): 230–38; F. Tong, "Primary Visual Cortex and Visual Awareness," *Nature Reviews Neuroscience* 4 (2003): 219–29.

40. A. Cowey, "The Blindsight Saga," *Experimental Brain Research* 200 (2010): 3–24; L. Weiskrantz, E. K. Warrington, M. D. Sanders, and J. Marshall, "Visual Capacity in the Hemianopic Field following a Restricted Cortical Ablation," *Brain* 97 (1974): 709–28.

41. T. N. Aflalo and M. S. A. Graziano, "Organization of the Macaque Extrastriate Visual Cortex Re-examined Using the Principle of Spatial Continuity of Function," *Journal of Neurophysiology* 105 (2011): 305–20; D. Felleman and D. Van Essen, "Distributed Hierarchical Processing in the Primate Visual Cortex," *Cerebral Cortex* 1 (1991): 1–47; M. A. Goodale and A. D. Milner, "Separate Visual Pathways for Perception and Action," *Trends in Neurosciences* 15 (1992): 20–25; L. G. Ungerleider and J. V. Haxby, "'What' and 'Where' in the Human Brain," *Current Opinion in Neurobiology* 4 (1994): 157–65.

42. S. Brown and E. Schafer, "An Investigation into the Functions of the Occipital and Temporal Lobes of the Monkey's Brain," *Philosophical Transactions of the Royal Society of London, B, Biological Sciences* 179 (1888): 303–27.

43. P. Broca, "Remarks on the Seat of the Faculty of Articulate Language, Followed by an Observation of Aphemia," *Bulletin de la Societe Anatomique de Paris* 6 (1861): 330–57, trans. G. von Bonin and republished in *Some Papers on the Cerebral Cortex*, ed. G. Von Bonin (Springfield, IL: Charles Thomas Publisher, 1960), 49–72; A. R. Damasio and N. Geschwind, "The Neural Basis of Language," *Annual Review of Neuroscience* 7 (1984): 127–47.

44. J. Zihl, D. von Cramon, and N. Mai, "Selective Disturbance of Movement Vision after Bilateral Brain Damage," *Brain* 106 (1983): 313–40.

45. S. E. Bouvier and S. A. Engel, "Behavioral Deficits and Cortical Damage Loci in Cerebral Achromatopsia," *Cerebral Cortex* 16 (2006):

183–91.

46. C. Gottesmann, “The Neurophysiology of Sleep and Waking: Intracerebral Connections, Functioning and Ascending Influences of the Medulla Oblongata,” *Progress in Neurobiology* 59 (1999): 1–54.

47. D. Chalmers, *The Conscious Mind* (Oxford, UK: Oxford University Press); P. Skokowski, “I, Zombie,” *Consciousness and Cognition* 11 (2002): 1–9; C. Tandy, “Are You (Almost) a Zombie? Conscious Thoughts about ‘Consciousness in the Universe’ by Hameroff and Penrose,” *Physics of Life Reviews* 11 (2014): 89–90.

48. W. R. Brain, “A Form of Visual Disorientation Resulting from Lesions of the Right Cerebral Hemisphere,” *Proceedings of the Royal Society of Medicine* 34 (1941): 771–76; M. Critchley, *The Parietal Lobes* (London: Hafner Press, 1953); G. Vallar, “Extrapersonal Visual Unilateral Spatial Neglect and Its Neuroanatomy,” *Neuroimage* 14 (2001): S52–S58.

49. K. M. Heilman and E. Valenstein “Mechanism Underlying Hemispatial Neglect,” *Annual Neurology* 5 (1972): 166–70; M. Kinsbourne, “A Model for the Mechanism of Unilateral Neglect of Space,” *Transactions of the American Neurological Association* 95 (1970): 143–46; M. M. Mesulam, “A Cortical Network for Directed Attention and Unilateral Neglect,” *Annual Neurology* 10 (1981): 309–25; S. M. Szczepanski, C. S. Konen, and S. Kastner, “Mechanisms of Spatial Attention Control in Frontal and Parietal Cortex,” *Journal of Neuroscience* 30 (2010): 148–60.

50. P. Chen and K. M. Goedert, “Clock Drawing in Spatial Neglect: A Comprehensive Analysis of Clock Perimeter, Placement, and Accuracy,” *Journal of Neuropsychology* 6 (2012): 270–89.

51. E. Bisiach and C. Luzzatti, “Unilateral Neglect of Representational Space,” *Cortex* 14 (1978): 129–33.

52. J. C. Marshall and P. W. Halligan, “Blindsight and Insight in Visuo-Spatial Neglect,” *Nature* 336 (1988): 766–67.

53. G. Vallar and D. Perani, “The Anatomy of Unilateral Neglect after Right-Hemisphere Stroke Lesions: A Clinical/CT-Scan Correlation Study in Man,” *Neuropsychologia* 24 (1986): 609–22.

54. M. A. Bruno, S. Majerus, M. Boly, A. Vanhaudenhuyse, C. Schnakers, O. Gosseries, P. Boveroux, et al., “Functional Neuroanatomy Underlying the Clinical Subcategorization of Minimally Conscious State Patients,” *Journal of Neurology* 259 (2012): 1087–98; S. Laureys, “The Neural Correlate of (Un) Awareness: Lessons from the Vegetative State,” *Trends in Cognitive Sciences* 9 (2005): 556–59; S. Laureys, S. Antoine, M. Boly, S. Elincx, M. E. Faymonville, J. Berré, B. Sadzot, et al., “Brain Function in the Vegetative State,” *Acta Neurologica Belgica* 102 (2002): 177–85; J. Leon-Carrion, U. Leon-Dominguez, L. Pollonini, M. H. Wu, R. E. Frye, M. R. Dominguez-Morales, and G. Zouridakis, “Synchronization between the Anterior and Posterior Cortex Determines Consciousness Level in Patients with Traumatic Brain Injury,” *Brain Research* 1476 (2012): 22–30; D. Roquet, J. R. Foucher, P. Froehlig, F. Renard, J. Pottecher, H. Besancenot, F.

Schneider, et al., “Resting-State Networks Distinguish Locked-In from Vegetative State Patients,” *Neuroimage: Clinical* 12 (2016): 16–22.

Глава 7. Трудная проблема и другие точки зрения на сознание

1. D. Chalmers, “Facing Up to the Problem of Consciousness,” *Journal of Consciousness Studies* 2 (1995): 200–219.
2. D. Chalmers, “The Meta-Problem of Consciousness,” *The Journal of Consciousness Studies* 25, nos. 9–10 (2018): 6–61.
3. I. A. Newton, “Letter of Mr. Isaac Newton, Professor of the Mathematicks in the University of Cambridge; Containing His New Theory about Light and Colors: Sent by the Author to the Publisher from Cambridge, Febr. 6. 1671/72; In Order to Be Communicated to the Royal Society,” *Philosophical Transactions Royal Society* 6 (1671): 3075–87.
4. F. Kammerer, “The Hardest Aspect of the Illusion Problem — And How to Solve It,” *Journal of Consciousness Studies* 23 (2016): 124–39; F. Kammerer, “Can You Believe It? Illusionism and the Illusion Meta-Problem,” *Philosophical Psychology* 31 (2018): 44–67.
5. S. Blackmore, “Delusions of Consciousness,” *Journal of Consciousness Studies* 23 (2016): 52–64; F. Crick, *The Astonishing Hypothesis: The Scientific Search for the Soul* (New York: Scribner, 1995); D. C. Dennett, *Consciousness Explained* (Boston: Back Bay Books, 1991); K. Frankish, “Illusionism as a Theory of Consciousness,” *Journal of Consciousness Studies* 23 (2016): 1–39; B. Hood, *The Self Illusion: How the Social Brain Creates Identity* (Oxford, UK: Oxford University Press, 2012); F. Kammerer, “The Hardest Aspect of the Illusion Problem — And How to Solve It,” *Journal of Consciousness Studies* 23 (2016): 124–39.
6. Утверждение, что сознание — иллюзия, нужная, чтобы сделать жизнь приятнее и продуктивнее, содержится в работе: N. Humphrey, *Soul Dust* (Princeton, NJ: Princeton University Press, 2011).
7. S. Glucksberg, *Understanding Figurative Language* (Oxford, UK: Oxford University Press, 2001).
8. V. S. Ramachandran and W. Hirstein, “The Perception of Phantom Limbs,” *Brain* 121 (1998): 1603–30; A. Woodhouse, “Phantom Limb Sensation,” *Clinical and Experimental Pharmacology and Physiology* 32 (2005): 132–34.
9. V. S. Ramachandran and W. Hirstein, “The Perception of Phantom Limbs,” *Brain* 121 (1998): 1603–30.
10. Y. Luo and T. A. Anderson, “Phantom Limb Pain: A Review,” *International Anesthesiology Clinics* 54 (2016): 121–39.
11. V. S. Ramachandran and W. Hirstein, “The Perception of Phantom Limbs,” *Brain* 121 (1998): 1603–30.
12. G. Vallar and R. Ronchi, “Somatoparaphrenia: A Body Delusion. A Review of the Neuropsychological Literature,” *Experimental Brain*

Research 192 (2009): 533–51.

13. O. Sacks, *The Man Who Mistook His Wife for a Hat* (New York: Touchstone, 1998), 56. Русский перевод: Сакс О. Человек, который принял жену за шляпу, и другие истории из врачебной практики. — АСТ, 2015.

14. M. Botvinick and J. D. Cohen, “Rubber Hand ‘Feels’ What Eye Sees,” *Nature* 391 (1998): 756.

15. M. S. A. Graziano, *The Spaces between Us: A Story of Neuroscience, Evolution, and Human Nature* (Oxford, UK: Oxford University Press, 2018).

16. O. Blanke and T. Metzinger, “Full-Body Illusions and Minimal Phenomenal Selfhood,” *Trends in Cognitive Sciences* 13 (2009): 7–13.

17. B. J. Baars, *A Cognitive Theory of Consciousness* (Cambridge, UK: Cambridge University Press, 1988).

18. S. Dehaene, *Consciousness and the Brain* (New York: Viking Press, 2014).

19. D. Dennett, *Sweet Dreams* (Cambridge, MA: MIT Press, 2005). Русский перевод: см. выше.

20. C. G. Gross, *Brain, Vision, Memory: Tales in the History of Neuroscience* (New York: Bradford Books, 1999).

21. H. Palsson and P. Edwards, *Seven Viking Romances* (Toronto, Canada: Penguin Books, 1985).

22. D. Rosenthal, *Consciousness and Mind* (Oxford, UK: Oxford University Press, 2006); см. также: R. L. Gennaro, *Consciousness and Self Consciousness: A Defense of the Higher Order Thought Theory of Consciousness* (Philadelphia: John Benjamin’s Publishing, 1996); H. Lau and D. Rosenthal, “Empirical Support for Higher-Order Theories of Consciousness,” *Trends in Cognitive Sciences* 15 (2011): 365–73.

23. P. Carruthers, “How We Know Our Own Minds: The Relationship between Mindreading and Metacognition,” *Behavioral and Brain Sciences* 32 (2009): 121–82; A. Pasquali, B. Timmermans, and A. Cleeremans, “Know Thyself: Metacognitive Networks and Measures of Consciousness,” *Cognition* 117 (2010): 182–90; D. M. Rosenthal, “Consciousness, Content, and Metacognitive Judgments,” *Consciousness and Cognition* 9 (2000): 203–14.

24. D. D. Hoffman, “The Interface Theory of Perception,” in *Object Categorization: Computer and Human Vision Perspectives*, ed. S. Dickinson, M. Tarr, A. Leonardis, and B. Schiele (New York: Cambridge University Press, 2009), 148–65.

25. P. Grimaldi, H. Lau, and M. A. Basso, “There Are Things That We Know That We Know, and There Are Things That We Do Not Know We Do Not Know: Confidence in Decision-Making,” *Neuroscience and Biobehavioral Reviews* 55 (2015): 88–97.

26. D. C. Dennett, *Consciousness Explained* (Boston: Back Bay Books, 1991).

27. S. J. Blackmore, “Consciousness in Meme Machines,” *Journal of Consciousness Studies* 10 (2003): 19–30.

28. W. James, *Principles of Psychology* (New York: Henry Holt & Co., 1890).

29. A. M. Turing, “On Computable Numbers, with an Application to the Entscheidungsproblem,” *Proceedings of the London Mathematical Society* S2–42 (1937): 230–65.

30. C. E. Shannon, “A Mathematical Theory of Communication,” *Bell System Technical Journal* 27 (1948): 379–423.

31. R. W. Kentridge, C. A. Heywood, and L. Weiskrantz, “Attention without Awareness in Blindsight,” *Proceedings: Biological Sciences* 266 (1999): 1805–11; R. W. Kentridge, C. A. Heywood, and L. Weiskrantz, “Spatial Attention Speeds Discrimination without Awareness in Blindsight,” *Neuropsychologia* 42 (2004): 831–35.

32. Здесь я привожу лишь небольшой ряд из множества блестящих работ, показывающих разделение осознания и внимания. Это явление заинтересовало многих исследователей и стало темой большого количества работ, поскольку оказалось одним из немногих экспериментальных результатов, напрямую связанных с сознанием, которые можно получить в контролируемых лабораторных условиях. U. Ansorge and M. Heumann, “Shifts of Visuospatial Attention to Invisible (Metacontrast-Masked) Singletons: Clues from Reaction Times and Event-Related Potentials,” *Advances in Cognitive Psychology* 2 (2006): 61–76; P. Hsieh, J. T. Colas, and N. Kanwisher, “Unconscious Pop-Out: Attentional Capture by Unseen Feature Singletons Only When Top-Down Attention Is Available,” *Psychological Science* 22 (2011): 1220–26; J. Ivanoff and R. M. Klein, “Orienting of Attention without Awareness Is Affected by Measurement-Induced Attentional Control Settings,” *Journal of Vision* 3 (2003): 32–40; Y. Jiang, P. Costello, F. Fang, M. Huang, and S. He, “A Gender- and Sexual Orientation-Dependent Spatial Attentional Effect of Invisible Images,” *Proceedings of the National Academy of Sciences USA* 103 (2006): 17048–52; R. W. Kentridge, T. C. Nijboer, and C. A. Heywood, “Attended but Unseen: Visual Attention Is Not Sufficient for Visual Awareness,” *Neuropsychologia* 46 (2008): 864–69; C. Koch and N. Tsuchiya, “Attention and Consciousness: Two Distinct Brain Processes,” *Trends in Cognitive Sciences* 11 (2007): 16–22; A. Lambert, N. Naikar, K. McLachlan, and V. Aitken, “A New Component of Visual Orienting: Implicit Effects of Peripheral Information and Subthreshold Cues on Covert Attention,” *Journal of Experimental Psychology, Human Perception and Performance* 25 (1999): 321–40; V. A. Lamme, “Separate Neural Definitions of Visual Consciousness and Visual Attention: A Case for Phenomenal Awareness,” *Neural Networks* 17 (2004): 861–72; Z. Lin and S. O. Murray, “More Power to the Unconscious: Conscious, but Not Unconscious, Exogenous Attention Requires Location Variation,” *Psychological Science* 26 (2015): 221–30; P. A. McCormick, “Orienting Attention without Awareness,” *Journal of Experimental Psychology, Human Perception and Performance* 23 (1997): 168–80; L. J. Norman,

C. A. Heywood, and R. W. Kentridge, “Object-Based Attention without Awareness,” *Psychological Science* 24 (2013): 836–43; Y. Tsushima, Y. Sasaki, and T. Watanabe, “Greater Disruption Due to Failure of Inhibitory Control on an Ambiguous Distractor,” *Science* 314 (2006): 1786–88; T. W. Webb, H. H. Kean, and M. S. A. Graziano, “Effects of Awareness on the Control of Attention,” *Journal of Cognitive Neuroscience* 28 (2016): 842–51; G. F. Woodman and S. J. Luck, “Dissociations among Attention, Perception, and Awareness during Object-Substitution Masking,” *Psychological Science* 14 (2003): 605–11.

33. Y. Tsushima, Y. Sasaki, and T. Watanabe, “Greater Disruption Due to Failure of Inhibitory Control on an Ambiguous Distractor,” *Science* 314 (2006): 1786–88; T. W. Webb, H. H. Kean, and M. S. A. Graziano, “Effects of Awareness on the Control of Attention,” *Journal of Cognitive Neuroscience* 28 (2016): 842–51.

34. T. W. Webb, H. H. Kean, and M. S. A. Graziano, “Effects of Awareness on the Control of Attention,” *Journal of Cognitive Neuroscience* 28 (2016): 842–51.

35. Практически невозможно упомянуть все великое множество теорий, включающих в себя представление о том, что сознание связано с интеграцией информации. Вот лишь немногие из них: B. J. Baars, *A Cognitive Theory of Consciousness* (Cambridge, UK: Cambridge University Press, 1988); A. B. Barrett, “An Integration of Integrated Information Theory with Fundamental Physics,” *Frontiers in Psychology* 5 (2014): 63; F. Crick and C. Koch, “Toward a Neurobiological Theory of Consciousness,” *Seminars in the Neurosciences* 2 (1990): 263–75; A. Damasio, *Self Comes to Mind: Constructing the Conscious Brain* (New York: Pantheon, 2015); S. Dehaene, *Consciousness and the Brain* (New York: Viking Press, 2014); G. M. Edelman, J. A. Gally, and B. J. Baars, “Biology of Consciousness,” *Frontiers in Psychology* 2 (2012): 4; A. K. Engel and W. Singer, “Temporal Binding and the Neural Correlates of Sensory Awareness,” *Trends in Cognitive Sciences* 5 (2011): 16–25; S. Grossberg, “The Link between Brain Learning, Attention, and Consciousness,” *Consciousness and Cognition* 8 (1999): 1–44; V. A. Lamme, “Towards a True Neural Stance on Consciousness,” *Trends in Cognitive Sciences* 10 (2006): 494–501; G. Tononi, M. Boly, M. Massimini, and C. Koch, “Integrated Information Theory: From Consciousness to Its Physical Substrate,” *Nature Reviews Neuroscience* 17 (2016): 450–61; C. Von der Malsburg, “The Coherence Definition of Consciousness,” in *Cognition, Computation, and Consciousness*, ed. M. Ito, Y. Miyashita, and E. Rolls (Oxford, UK: Oxford University Press, 1997), 193–204; L. M. Ward, “The Thalamic Dynamic Core Theory of Conscious Experience,” *Consciousness and Cognition* 20 (2011): 464–86.

36. G. Tononi, *Phi: A Voyage from the Brain to the Soul* (New York: Pantheon, 2012).

37. K. Koffka, *Principles of Gestalt Psychology* (New York: Harcourt, Brace, 1935).

38. B. E. Stein and M. A. Meredith, *The Merging of the Senses* (Cambridge, MA: MIT Press, 1993).

39. Синдром Балинта, вызванный повреждением частей теменной доли, может быть примером распада единого чувственного мира при нарушении пространственной информации. H. Udesen and A. L. Madsen, “Balint’s Syndrome — Visual Disorientation,” *Ugeskrift for Laeger* 154 (1992): 1492–94.

Глава 8. Машины, обладающие сознанием

1. M. White, *Isaac Newton: The Last Sorcerer* (New York: Basic Books, 1999).

2. I. Aleksander, *Impossible Minds: My Neurons, My Consciousness* (Singapore: World Scientific, 2014); B. J. Baars and S. Franklin, “Consciousness Is Computational: The LIDA Model of Global Workspace Theory,” *International Journal of Machine Consciousness* 1 (2009): 23–32; A. Chella and R. Manzotti, “Machine Consciousness: A Manifesto for Robotics,” *International Journal of Machine Consciousness* 1 (2009): 33–51; L. A. Coward and R. Sun, “Criteria for an Effective Theory of Consciousness and Some Preliminary Attempts,” *Consciousness and Cognition* 13 (2004): 268–301; S. Franklin, “IDA: A Conscious Artefact,” in *Machine Consciousness*, ed. O. Holland (Exeter, UK: Imprint Academic, 2003); P. Haikonen, *Consciousness and Robot Sentience* (Singapore: World Scientific, 2012); O. Holland and R. Goodman, “Robots with Internal Models: A Route to Machine Consciousness?” *Journal of Consciousness Studies* 10 (2003): 77–109; N. Marupaka, L. Lyer, and A. Minai, “Connectivity and Thought: The Influence of Semantic Network Structure in a Neurodynamical Model of Thinking,” *Neural Networks* 32 (2012): 147–58; D. Rudrauf, D. Bennequin, I. Granic, G. Landini, K. Friston, and K. Williford, “A Mathematical Model of Embodied Consciousness,” *Journal of Theoretical Biology* 428 (2017): 106–31; M. Shanahan, “A Cognitive Architecture That Combines Internal Simulation with a Global Workspace,” *Consciousness and Cognition* 15 (2006): 443–49.

3. A. M. Turing, “Computing Machinery and Intelligence,” *Mind* 59 (1950): 433–60.

4. S. Baron-Cohen, A. M. Leslie, and U. Frith, “Does the Autistic Child Have a ‘Theory of Mind?’” *Cognition* 21 (1985): 37–46; H. Wimmer and J. Perner, “Beliefs about Beliefs: Representation and Constraining Function of Wrong Beliefs in Young Children’s Understanding of Deception,” *Cognition* 13 (1983): 103–28.

5. N. S. Clayton, “Ways of Thinking: From Crows to Children and Back Again,” *Quarterly Journal of Experimental Psychology* 68 (2015): 209–41; C. Krupenye, F. Kano, S. Hirata, J. Call, and M. Tomasello, “Great Apes Anticipate That Other Individuals Will Act According to False Beliefs,” *Science* 354 (2016): 110–14; H. M. Wellman, D. Cross, and J. Watson, “Meta-Analysis of Theory-of-Mind Development: The Truth about False Belief,” *Child Development* 72 (2001): 655–84.

6. M. A. Lebedev and M. A. Nicolelis, “Brain-Machine Interfaces: From Basic Science to Neuroprostheses and Neurorehabilitation,” *Physiological Review* 97 (2017): 767–837.

7. J. V. Haxby, M. I. Gobbini, M. L. Furey, A. Ishai, J. L. Schouten, and P. Pietrini, “Distributed and Overlapping Representations of Faces and Objects in Ventral Temporal Cortex,” *Science* 293 (2001): 2425–30.

8. Внимание так часто встраивали в искусственные устройства, что я в состоянии привести лишь весьма неполный список литературы: H. Adeli, F. Vitu, and G. F. Zelinsky, “A Model of the Superior Colliculus Predicts Fixation Locations during Scene Viewing and Visual Search,” *Journal of Neuroscience* 37 (2017): 1453–67; A. Borji and L. Itti, “State-of-the-Art in Visual Attention Modeling,” *IEEE Transactions on Pattern Analysis and Machine Intelligence* 35 (2013): 185–207; G. Deco and E. T. Rolls, “A Neurodynamical Cortical Model of Visual Attention and Invariant Object Recognition,” *Vision Research* 44 (2004): 621–42; Y. Fang, C. Zhang, J. Li, J. Lei, M. Perreira Da Silva, and P. Le Callet, “Visual Attention Modeling for Stereoscopic Video: A Benchmark and Computational Model,” *IEEE Transactions on Image Processing* 26 (2017): 4684–96; S. Goferman, L. Zelnikmanor, and A. Tal, “Context-Aware Saliency Detection,” *IEEE Transactions on Pattern Analysis and Machine Intelligence* 34 (2012): 1915–26; C. Guo and L. Zhang, “A Novel Multi-Resolution Spatiotemporal Saliency Detection Model and Its Applications in Image and Video Compression,” *IEEE Transactions on Image Processing* 19 (2010): 185–98; L. Itti, C. Koch, and E. Niebur, “A Model of Saliency-Based Visual Attention for Rapid Scene Analysis,” *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20 (1988): 1254–59; O. Le Meur, P. Le Callet, and D. Barba, “A Coherent Computational Approach to Model the Bottom-Up Visual Attention,” *IEEE Transactions on Pattern Analysis and Machine Intelligence* 28 (2006): 802–17; R. J. Lin and W. S. Lin, “A Computational Visual Saliency Model Based on Statistics and Machine Learning,” *Journal of Vision* 14 (2014): 1; T. Miconi and R. VanRullen, “A Feedback Model of Attention Explains the Diverse Effects of Attention on Neural Firing Rates and Receptive Field Structure,” *PLoS Computational Biology* 12 (2016): e1004770; J. H. Reynolds and D. J. Heeger, “The Normalization Model of Attention,” *Neuron* 61 (2009): 168–85; P. Schwedhelm, B. S. Krishna, and S. Treue, “An Extended Normalization Model of Attention Accounts for Feature-Based Attentional Enhancement of Both Response and Coherence Gain,” *PLoS Computational Biology* 12 (2016): e1005225; M. A. Schwemmer, S. F. Feng, P. J. Holmes, J. Gottlieb, and J. D. Cohen, “A Multi-Area Stochastic Model for a Covert Visual Search Task,” *PLoS One* 10 (2015): e0136097; S. Vossel, C. Mathys, K. E. Stephan, and K. J. Friston, “Cortical Coupling Reflects Bayesian Belief Updating in the Deployment of Spatial Attention,” *Journal of Neuroscience* 35 (2015): 11532–42; A. L. White, M. Rolfs, and M. Carrasco, “Stimulus Competition Mediates the Joint Effects of Spatial and Feature-Based Attention,” *Journal of Vision* 15 (2015). doi: 10.1167/15.14.7; P. Zhang, T. Zhuo, W. Huang, K. Chen, and M. Kankanhalli, “Online Object Tracking Based on CNN with Spatial-

Temporal Saliency Guided Sampling,” *Neurocomputing* 257 (2017): 115–27.

9. E. van den Boogaard, J. Treur, and M. Turpijn, “A Neurologically Inspired Neural Network Model for Graziano’s Attention Schema Theory for Consciousness,” *International Work Conference on the Interplay between Natural and Artificial Computation: Natural and Artificial Computation for Biomedicine and Neuroscience* Part 1 (2017): 10–21.

10. M. M. Chun, J. D. Golomb, and N. B. Turk-Browne, “A Taxonomy of External and Internal Attention,” *Annual Review of Psychology* 62 (2011): 73–101.

11. J. Ledoux, *The Emotional Brain: The Mysterious Underpinnings of Emotional Life* (New York: Simon & Schuster, 1998).

12. W. R. Hess, *Functional Organization of the Diencephalons* (New York: Grune and Stratton, 1957).

13. B. G. Hoebel, “Neuroscience and Appetitive Behavior Research: 25 Years,” *Appetite* 29 (1997): 119–33; T. V. Sewards and M. A. Sewards, “Representations of Motivational Drives in Mesial Cortex, Medial Thalamus, Hypothalamus and Midbrain,” *Brain Research Bulletin* 61 (2003): 25–49; A. Venkatraman, B. L. Edlow, and M. H. Immordino-Yang, “The Brainstem in Emotion: A Review,” *Frontiers in Neuroanatomy* 11 (2017): 15.

14. J. LeDoux, “The Amygdala,” *Current Biology* 17 (2007): R868–R874; P. J. Walen and E. A. Phelps, *The Human Amygdala* (New York: Guilford Press, 2009).

15. E. T. Rolls and F. Grabenhorst, “The Orbitofrontal Cortex and Beyond: From Affect to Decision-Making,” *Progress in Neurobiology* 86 (2008): 216–44.

16. M. Tamietto and B. de Gelder, “Neural Bases of the Non-Conscious Perception of Emotional Signals,” *Nature Reviews Neuroscience* 11 (2010): 697–709; P. Winkielman and K. C. Berridge, “Unconscious Emotion,” *Current Directions in Psychological Science* 13 (2004): 120–23.

17. J. E. LeDoux and R. Brown, “A Higher-Order Theory of Emotional Consciousness,” *Proceedings of the National Academy of Sciences, USA* 114 (2017): E2016–E2025.

18. W. Cannon, “The James-Lange Theory of Emotions: A Critical Examination and an Alternative Theory,” *The American Journal of Psychology* 39 (1927): 106–24.

19. D. G. Dutton and A. P. Aaron, “Some Evidence for Heightened Sexual Attraction under Conditions of High Anxiety,” *Journal of Personality and Social Psychology* 30 (1974): 510–17.

20. M. E. Moran, “The da Vinci Robot,” *Journal of Endourology* 20 (2006): 986–90.

21. I. Asimov, *The Bicentennial Man* (New York: Ballantine Books, 1976). Русский перевод И. Гуровой. Двухсотлетний человек. В кн.:

Азимов А. Мечты роботов: Фантастические произведения. — М.: Эксмо, 2009. — С. 262.

22. P. K. Dick, *Do Androids Dream of Electric Sheep?* (New York: Doubleday, 1968). Русский перевод Н. Романецкого. Мечтают ли андроиды об электрических овцах? В кн.: Дик Ф. К. В ожидании прошлого: Сб. — М.: ООО “Издательство АСТ”: ЗАО НПП “Ермак”, 2004.

23. D. Levy, “The Ethical Treatment of Artificially Conscious Robots,” *International Journal of Social Robotics* 1 (1929): 209–16.

24. B. Hood, *The Self Illusion: How the Social Brain Creates Identity* (Oxford, UK: Oxford University Press, 2012); F. Podschwadek, “Do Androids Dream of Normative Endorsement? On the Fallibility of Artificial Moral Agents,” *Artificial Intelligence and Law* 25 (2017): 325–39; J. Sullins, “Artificial Phronesis and the Social Robot,” *Frontiers in Artificial Intelligence and Applications* 290 (2016): 37–39.

Глава 9. Перенос личности на искусственные носители

1. R. Blackford and D. Broderick, eds., *Intelligence Unbound: The Future of Uploads and Machine Minds* (Hoboken, NJ: Wiley Blackwell, 2014); C. Eliasmith, T. C. Stewart, X. Choo, T. Bekolay, T. DeWolf, Y. Tang, and D. Rasmussen, “A Large-Scale Model of the Functioning Brain,” *Science* 338 (2012): 1202–05; D. Eth, J.-C. Foust, and B. Whale, “The Prospects of Whole Brain Emulation within the Next Half-Century,” *Journal of Artificial General Intelligence* 4 (2013): 130–52; R. A. Koene, “Feasible Mind Uploading,” in *Intelligence Unbound: The Future of Uploaded and Machine Minds*, ed. R. Blackford and D. Broderick (Hoboken, NJ: Wiley-Blackwell, 2014); R. Kurzweil, *The Singularity Is Near: When Humans Transcend Biology* (New York: Penguin Books, 2006); H. Markram, E. Muller, S. Ramaswamy, M. Reimann, M. Abdellah, C. A. Sanchez, A. Ailamaki, et al., “Reconstruction and Simulation of Neocortical Microcircuitry,” *Cell* 163 (2015): 456–92; H. Moravec, *Mind Children: The Future of Robot and Human Intelligence* (Cambridge, MA: Harvard University Press, 1988).

2. S. Herculano-Houzel, “The Human Brain in Numbers: A Linearly Scaled-Up Primate Brain,” *Frontiers in Human Neuroscience* 3 (2009). doi: 10.3389/neuro.09.031.2009

3. C. S. Sherrington, “Santiago Ramón y Cajal 1852–1934,” *Biographical Memoirs of Fellows of the Royal Society* 1 (1935): 424–41.

4. S. R. Cajal, J. DeFelipe, and E. G. Jones, *Cajal on the Cerebral Cortex: An Annotated Translation of the Complete Writings* (Oxford, UK: Oxford University Press, 1988).

5. D. E. Rumelhart and J. McClelland, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* (Cambridge, MA: MIT Press, 1986); J. Schmidhuber, “Deep Learning in Neural Networks: An Overview,” *Neural Networks* 61 (2015): 85–117.

6. D. M. Barch, “Resting-State Functional Connectivity in the Human Connectome Project: Current Status and Relevance to Understanding

Psychopathology," *Harvard Review of Psychiatry* 25 (2017): 209–17; D. D. Bock, W. C. Lee, A. M. Kerlin, M. L. Andermann, G. Hood, A. W. Wetzel, S. Yurgenson, et al., "Network Anatomy and *In Vivo* Physiology of Visual Cortical Neurons," *Nature* 471 (2011): 177–82; G. Gong, Y. He, L. Concha, C. Lebel, D. W. Gross, A. C. Evans, and C. Beaulieu, "Mapping Anatomical Connectivity Patterns of Human Cerebral Cortex Using *In Vivo* Diffusion Tensor Imaging Tractography," *Cerebral Cortex* 19 (2009): 524–36; P. Hagmann, L. Cammoun, X. Gigandet, R. Meuli, C. J. Honey, V. J. Wedeen, and O. Sporns, "Mapping the Structural Core of Human Cerebral Cortex," *PLoS Biol* 6 (2008): e159; P. Hagmann, M. Kurant, X. Gigandet, P. Thiran, V. J. Wedeen, R. Meuli, and J.-P. Thiran, "Mapping Human Whole-Brain Structural Networks with Diffusion MRI," *PLoS One* 2 (2007): e597; M. Helmstaedter, K. L. Briggman, S. C. Turaga, V. Jain, H. S. Seung, and W. Denk, "Connectomic Reconstruction of the Inner Plexiform Layer in the Mouse Retina," *Nature* 500 (2013): 168–74; O. Sporns, G. Tononi, and R. Kötter, "The Human Connectome: A Structural Description of the Human Brain," *PLoS Computational Biology* 1 (2005): e42; L. R. Varshney, B. L. Chen, E. Paniagua, D. H. Hall, and D. B. Chklovskii, "Structural Properties of the *Caenorhabditis elegans* Neuronal Network," *PLoS Computational Biology* 7 (2011): e1001066; Z. Zheng, J. S. Lauritzen, E. Perlman, C. G. Robinson, M. Nichols, D. Milkie, O. Torrens, et al., "A Complete Electron Microscopy Volume of the Brain of Adult *Drosophila melanogaster*," *Cell* 174 (2018): 730–43.

7. L. R. Varshney, B. L. Chen, E. Paniagua, D. H. Hall, and D. B. Chklovskii, "Structural Properties of the *Caenorhabditis elegans* Neuronal Network," *PLoS Computational Biology* 7 (2011): e1001066; Z. Zheng, J. S. Lauritzen, E. Perlman, C. G. Robinson, M. Nichols, D. Milkie, O. Torrens, et al., "A Complete Electron Microscopy Volume of the Brain of Adult *Drosophila melanogaster*," *Cell* 174 (2018): 730–43.

8. D. D. Bock, W. C. Lee, A. M. Kerlin, M. L. Andermann, G. Hood, A. W. Wetzel, S. Yurgenson, et al., "Network Anatomy and *In Vivo* Physiology of Visual Cortical Neurons," *Nature* 471 (2011): 177–82.

9. D. M. Barch, "Resting-State Functional Connectivity in the Human Connectome Project: Current Status and Relevance to Understanding Psychopathology," *Harvard Review of Psychiatry* 25 (2017): 209–17; G. Gong, Y. He, L. Concha, C. Lebel, D. W. Gross, A. C. Evans, and C. Beaulieu, "Mapping Anatomical Connectivity Patterns of Human Cerebral Cortex Using *In Vivo* Diffusion Tensor Imaging Tractography," *Cerebral Cortex* 19 (2009): 524–36; P. Hagmann, L. Cammoun, X. Gigandet, R. Meuli, C. J. Honey, V. J. Wedeen, and O. Sporns, "Mapping the Structural Core of Human Cerebral Cortex," *PLoS Biol* 6 (2008): e159; P. Hagmann, M. Kurant, X. Gigandet, P. Thiran, V. J. Wedeen, R. Meuli, and J.-P. Thiran, "Mapping Human Whole-Brain Structural Networks with Diffusion MRI," *PLoS One* 2 (2007): e597; O. Sporns, G. Tononi, and R. Kötter, "The Human Connectome: A Structural Description of the Human Brain," *PLoS Computational Biology* 1 (2005): e42.

10. S. Herculano-Houzel, "The Human Brain in Numbers: A Linearly Scaled-Up Primate Brain," *Frontiers in Human Neuroscience* 3 (2009).

doi: 10.3389/neuro.09.031.2009

11. N. A. O'Rourke, N. C. Weiler, K. D. Micheva, and S. J. Smith, "Deep Molecular Diversity of Mammalian Synapses: Why It Matters and How to Measure It," *Nature Reviews Neuroscience* 13 (2012): 365–79; V. Pickel and M. Segal, *The Synapse: Structure and Function* (New York: Academic Press, 2014).

12. B. A. Barres, B. Stevens, and M. R. Freeman, *Glia* (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 2014).

13. A. Einstein, *The Collected Papers of Albert Einstein: Vol. 7: The Berlin Years: Writings, 1918–1921*, trans. A. Engel (Princeton, NJ: Princeton University Press, 2002).

14. B. P. Abbott et al. (LIGO Scientific Collaboration and Virgo Collaboration), "Observation of Gravitational Waves from a Binary Black Hole Merger," *Physical Review Letters* 116 (2016): 061102.

15. M. L. Cappuccio, "Mind-Upload. The Ultimate Challenge to the Embodied Mind Theory," *Phenomenology and the Cognitive Sciences* 16 (2017): 425–48; M. Wheeler, "Cognition in Context: Phenomenology, Situated Robotics and the Frame Problem," *International Journal of Philosophical Studies* 16 (2008): 323–49.

16. O. Blanke and T. Metzinger, "Full-Body Illusions and Minimal Phenomenal Selfhood," *Trends in Cognitive Sciences* 13 (2009): 7–13; M. S. A. Graziano and M. M. Botvinick, "How the Brain Represents the Body: Insights from Neurophysiology and Psychology," in *Common Mechanisms in Perception and Action: Attention and Performance XIX*, ed. W. Prinz and B. Hommel (Oxford, UK: Oxford University Press, 2002), 136–57; C. Lopez, "Making Sense of the Body: The Role of Vestibular Signals," *Multisensory Research* 28 (2015): 525–57; A. Serino, A. Alsmith, M. Costantini, A. Mandrigin, A. Tajadura-Jimenez, and C. Lopez, "Bodily Ownership and Self-Location: Components of Bodily Self-Consciousness," *Consciousness and Cognition* 22 (2013): 1239–52.

17. Статья, посвященная имитации руки, не была опубликована. Обзор моих исследований контроля движения см. в работе: M. S. A. Graziano, *The Intelligent Movement Machine* (Oxford, UK: Oxford University Press, 2008).

18. T. D. Bancroft, "Ethical Aspects of Computational Neuroscience," *Neuroethics* 6 (2013): 415–18; P. Eckersley and A. Sandberg, "Is Brain Emulation Dangerous?" *Journal of Artificial General Intelligence* 4 (2013): 170–94; K. Muzyka, "The Outline of Personhood Law Regarding Artificial Intelligences and Emulated Human Entities," *Journal of Artificial General Intelligence* 4 (2013): 164–69.

19. D. Sheils, "Toward a Unified Theory of Ancestor Worship: A Cross-Cultural Study," *Social Forces* 54 (1975): 427–40.

20. B. B. Powell, *Writing: Theory and History of the Technology of Civilization* (Oxford, UK: Blackwell Press, 2009).

21. J. P. Mallory and D. Q. Adams, *The Oxford Introduction to Proto-Indo-European and the Proto-Indo-European World* (Oxford, UK: Oxford

University Press, 2006).

22. G. Santayana, *Reason in Common Sense* (New York: Dover, 1980).

Приложение. Как сконструировать сознательное восприятие зрительной информации

1. R. Klette, *Concise Computer Vision* (New York: Springer, 2014); L. G. Shapiro and G. C. Stockman, *Computer Vision* (Upper Saddle River, NJ: Prentice Hall, 2001); M. Sonka, V. Hlavac, and R. Boyle, *Image Processing, Analysis, and Machine Vision* (Stamford, CT: Cengage Learning, 2008).

2. P. M. Merikle, D. Smilek, and J. D. Eastwood, "Perception without Awareness: Perspectives from Cognitive Psychology," *Cognition* 79 (2001): 115–34; R. Szczepanowski and L. Pessoa, "Fear Perception: Can Objective and Subjective Awareness Measures Be Dissociated?" *Journal of Vision* 10 (2007): 1–17.

3. M. Tegmark, "Consciousness as a State of Matter," *arXiv* (2014): 1401.1219; G. Tononi, M. Boly, M. Massimini, and C. Koch, "Integrated Information Theory: From Consciousness to Its Physical Substrate," *Nature Reviews Neuroscience* 17 (2016): 450–61.

4. M. S. A. Graziano and M. M. Botvinick, "How the Brain Represents the Body: Insights from Neurophysiology and Psychology," in *Common Mechanisms in Perception and Action: Attention and Performance XIX*, ed. W. Prinz and B. Hommel (Oxford, UK: Oxford University Press, 2002), 136–57; N. Holmes and C. Spence, "The Body Schema and the Multisensory Representation (s) of Personal Space," *Cognitive Processing* 5 (2004): 94–105; F. de Vignemont, *Mind the Body: An Exploration of Bodily Self-Awareness* (Oxford, UK: Oxford University Press, 2018).

5. S. Bluck and T. Habermas, "The Life Story Schema," *Motivation and Emotion* 24 (2000): 121–47; M. A. Conway and C. W. Pleydell-Pearce, "The Construction of Autobiographical Memories in the Self-Memory System," *Psychological Review* 107 (2000): 261–88; M. A. Conway, J. A. Singer, and A. Tagini, "The Self and Autobiographical Memory: Correspondence and Coherence," *Social Cognition* 22 (2004): 491–529.

6. O. Blanke, "Multisensory Brain Mechanisms of Bodily Self-Consciousness," *Nature Reviews Neuroscience* 13 (2012): 556–71; O. Blanke and T. Metzinger, "Full-Body Illusions and Minimal Phenomenal Selfhood," *Trends in Cognitive Sciences* 13 (2009): 7–13; C. Preston, B. J. Kuper-Smith, and H. H. Ehrsson, "Owning the Body in the Mirror: The Effect of Visual Perspective and Mirror View on the Full-Body Illusion," *Scientific Reports* 5 (2015): 18345.

7. M. S. Gazzaniga, *The Bisected Brain* (New York: Appleton Century Crofts, 1970); R. E. Nisbett and T. D. Wilson, "Telling More Than We Can Know — Verbal Reports on Mental Processes," *Psychological Review* 84 (1977): 231–59.

8. D. M. Beck and S. Kastner, "Top-Down and Bottom-Up Mechanisms in Biasing Competition in the Human Brain," *Vision Research* 49 (2009): 1154–65; R. Desimone and J. Duncan, "Neural Mechanisms of Selective Visual Attention," *Annual Review of Neuroscience* 18 (1995): 193–222.

9. G. Deco and E. T. Rolls, "Neurodynamics of Biased Competition and Cooperation for Attention: A Model with Spiking Neurons," *Journal of Neurophysiology* 94 (2005): 295–313; L. Layon and S. L. Denham, "A Biased Competition Computational Model of Spatial and Object-Based Attention Mediating Active Visual Search," *Neurocomputing* 58 (2004): 655–62; J. Reynolds and D. Heeger, "The Normalization Model of Attention," *Neuron* 61 (2009): 168–85; J. K. Tsotsos, *A Computational Perspective on Visual Attention* (Cambridge, MA: MIT Press, 2011).

[1] Elephant in the Room (*англ.*) — идиоматическое выражение, используемое для характеристики чего-то настолько бросающегося в глаза, что не заметить его сложно, однако наблюдатели либо и в самом деле не видят проблемы, либо по каким-то причинам предпочитают не обращать на нее внимания. Ср. с выражением “Слона-то я и не приметил”. — *Прим. ред.*

[2] Здесь авторская неточность. Роджер Сперри проводил подобные эксперименты в начале 1940-х гг. Работа 1943 г., на которую ссылается автор в Примечаниях, посвящена исследованию зрения тритонов без регенерации нерва. Упомянутый выше эксперимент был описан в работе 1944 г. “Optic nerve regeneration with return of vision in anurans”, опубликованной в *Journal of neurophysiology*. Полное библиографическое описание статьи см. в Примечаниях на с. 224. — *Прим. науч. ред.*

[3] Это один из немногих терминов, которые принято передавать латиницей и в русскоязычных научных текстах. — *Прим. науч. ред.*

[4] Авторская неточность. В басне “Ворона и кувшин” речь идет о том, что вороне удалось просто напиться воды. Но в статье, на которую ссылается автор, действительно описывается эксперимент с поднятием кусочка пищи. — *Прим. ред.*

[5] На русском языке есть ряд очень интересных книг на эту тему: Зорина З. А., Полетаева И. И. Зоопсихология. Элементарное мышление животных. — М.: Аспект-Пресс, 2010; Зорина З. А., Полетаева И. И. Я познаю мир. Поведение животных: Детская энциклопедия. — М.: АСТ: Астрель, 2000; Зорина З. А., Полетаева И. И., Резникова Ж. И. Основы этологии и

генетики поведения. — М.: Изд-во МГУ, 2013. — *Прим. науч. ред.*

[6] Финальная игра НФЛ. Речь идет об американском футболе. — *Прим. ред.*

[7] Это не совсем точно — не только коры. Общеизвестна гиппокампальная эпилепсия. Об участии других структур см. работу: Nowack, W. J., & Theodoridis, G. C. (1991). The thalamocortical contribution to epilepsy. *Bulletin of mathematical biology, 53* (4), 505–523. doi: 10.1007/BF02458626. — *Прим. науч. ред.*

[8] Автор здесь и далее использует именно такой порядок: теменно-лобные, хотя принято говорить о лобно-теменных сетях. — *Прим. науч. ред.*

[9] В русскоязычной научной психологической литературе термин “Theory of Mind” принято переводить как “модель психического”. — *Прим. науч. ред.*

[10] Сам не знаю что (*фр.*). — *Прим. пер.*

[11] Новые технологии в изучении функционирования мозга помогают понять, что этот метод не обладает никакими достоинствами помимо того, что он единственный был доступен для первых исследователей. — *Прим. науч. ред.*

[12] В ее основе лежит свойство нейропластичности мозга. — *Прим. науч. ред.*

[13] There’s no there there (*англ.*) — высказывание американской писательницы Гертруды Стайн о месте, где находился дом, в котором она выросла (Окленд, Калифорния). — *Прим. пер.*

[14] Цит. по: Оливер Сакс. Нога как точка опоры / Пер. А. В. Александровой. — М.: АСТ, 2014. — *Прим.*

пер.

[15] Фамилия этого исследователя произносится именно так, хотя есть переводы его работ на русский язык, где фамилия ученого передана как Деан. — *Прим. науч. ред.*

[16] Цит. по изданию: Дормашев Ю. Б., Романов В. Я. Психология внимания. — М.: Тривола, 1995. С. 283. — *Прим. ред.*

[17] Под квалиа понимается феноменальное качество осознаваемого переживания. — *Прим. науч. ред.*

[18] В данном случае имеется в виду функциональная магнитно-резонансная томография. — *Прим. науч. ред.*

[19] На официальном языке название премии звучит как Нобелевская премия по физиологии или медицине. — *Прим. ред.*

[20] Это упрощенное представление ныне претерпевает изменения: все большее внимание уделяется активному участию глиальных клеток, объемной передаче сигналов вне синаптических окончаний и другим аспектам. Автор пишет об этом дальше. — *Прим. науч. ред.*

[21] Недавние подсчеты выявили, что глиальных клеток всего в полтора раза больше или вообще столько же, сколько нейронов. — *Прим. науч. ред.*

[22] Emeritus — почетное звание профессора, вышедшего на пенсию, но не прекратившего преподавательской деятельности. — *Прим. ред.*

[23] Обозначение скорости света, равной 299 792 458 ± 1,2 м/с. — *Прим. ред.*

Книга издана при поддержке Политехнического музея и Фонда развития Политехнического музея.

Переводчик *Анна Петрова*

Научный редактор *Ольга Сварник, канд. психол. наук*

Редактор *Анастасия Ростоцкая*

Оформление серии *Андрея Бондаренко и Дмитрия Черногаева*

Издатель *П. Подкосов*

Руководитель проекта *И. Серёгина*

Корректоры *И. Астапкина, О. Петрова*

Компьютерная верстка *М. Зинуллин*

Дизайн обложки *Д. Черногаев*

Грациано М.

Наука сознания. Современная теория субъективного опыта / Майкл Грациано; Пер. с англ. А. Петровой. — М.: Альпина нон-фикшн, 2021.

ISBN 978-5-0013-9384-9